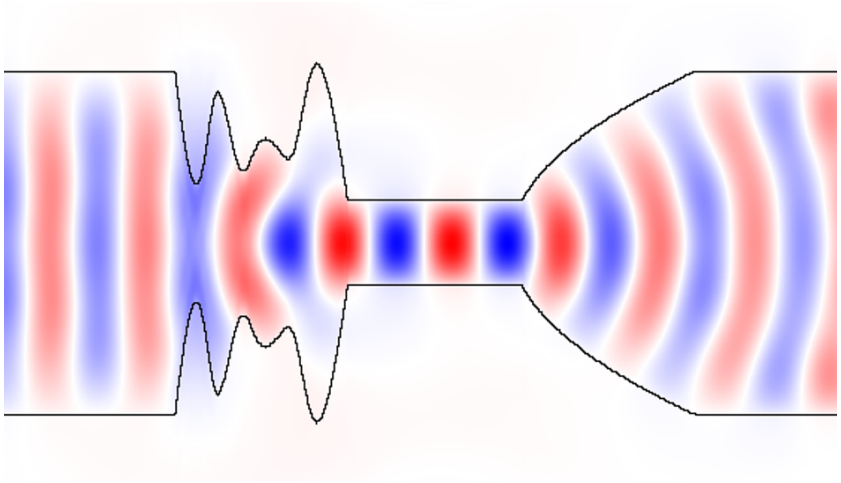
Compacte vlakke-golfgeleiderkoppelingen in silicium-op-isolator

Compact Planar Waveguide Spot-Size Converters in Silicon-on-Insulator

---

Bert Luyssaert



Promotoren: prof. dr. ir. R. Baets en prof. dr. ir. P. Bienstman
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Toegepaste Wetenschappen: Natuurkunde

Universiteit Gent
Faculteit Ingenieurswetenschappen
Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. P. Lagasse
Academiejaar: 2004-2005

Promotoren:

Prof. dr. ir. R. Baets	Universiteit Gent, INTEC
Prof. dr. ir. P. Bienstman	Universiteit Gent, INTEC

Examencommissie:

Prof. dr. ir. D. De Zutter (voorzitter)	Universiteit Gent, INTEC
Prof. dr. ir. D. Van Thourhout (secretaris)	Universiteit Gent, INTEC
Prof. dr. ir. O.J.F. Martin	EPFL Lausanne
Dr. X.J.M. Leijtens	TU/e Eindhoven
Prof. dr. ir. K. Neyts	Universiteit Gent, ELIS
Prof. dr. ir. R. Baets	Universiteit Gent, INTEC
Prof. dr. ir. P. Bienstman	Universiteit Gent, INTEC
Dr. ir. W. Bogaerts	Universiteit Gent, INTEC

Universiteit Gent
Faculteit Ingenieurswetenschappen

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

Tel.: +32-9-264.33.24
Fax: +32-9-264.35.93
http://www.intec.ugent.be

Dit werk kwam tot stand in het kader van een specialisatiebeurs van het IWT - Instituut voor de Aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen.

This work was carried out in the context of a specialization grant from the Flemish Institute for the Industrial Advancement of Scientific and Technological Research (IWT).

Every change is a menace to stability. That's another reason why we're so scary of applying new inventions. Every discovery in pure science is potentially subversive, even science must sometimes be treated as a possible enemy. Yes, even science.
[...]
Yes, Mustapha Mond was saying, that's another item in the cost of stability. It isn't only art that's incompatible with happiness, it's also science. Science is dangerous, we have to keep it most carefully chained and muzzled.

Brave New World, *Aldous Huxley*

There will be no art, no literature, no science. When we are omnipotent, we shall have no more need of science. There will be no distinction between beauty and ugliness. There will be no curiosity, no enjoyment of the process of life. All competing pleasures will be destroyed.
[...]
In Newspeak there is no word for 'Science'.

1984, *George Orwell*

Voorwoord - Preface

Doctoreren is best wel een leuke job. Niet enkel mag je 's morgens langer blijven liggen dan je partner, je ziet ook nog een stukje van de wereld. Zo mochten Delft, Eindhoven, Amsterdam, Lecce (I), Los Angeles, Ascona (S), Kopenhagen, Malmö, Grasmere (UK), Glasgow, Parijs, Brussel, Stockholm en - bijna - Warschau mij verwelkomen op een wetenschapstrip verbonden aan dit doctoraat.
Maar niet alleen je blik op de wereld verruimt, ook je garderobe wordt grondig aangepakt. Een IWT-verdediging dwong me tot de aankoop van een eerste kostuum en een symposium in Groot-Britannië met op de uitnodiging *jacket and tie* voegde er uiteindelijk ook een das aan toe. Een nachtelijke workshop in dasknopen op Wims trouwfeest vervolledigde dit onderdeel.
Toch is het vooral het wetenschappelijke gedeelte dat een doctoraat een leuke job maakt: het gevoel iets nieuws te ontdekken, er eindeloos over kunnen discussiëren, dit ideetje eens proberen, falen, dat ideetje eens proberen, . . .
In dit kader zou ik Roel willen bedanken voor de opportuniteit om onderzoek te verrichten in een groep die tegelijk competitief meedraait en waar iedereen toch zijn eigen vrijheid mag behouden, ongetwijfeld een zware evenwichtsoefening.
Voor de meer alledaagse beslommeringen zou ik een hele reeks collega's willen bedanken: op de eerste plaats Peter voor de ontelbare 'vraagjes' rond CAMFR en Python, Dirk voor de SEM- en meetlessen, Wim voor de PC tips and tricks, de gehele ploeg van de 39 voor de traditionele koffiepauzes en het occasionele H_2O-gevecht, en in het bijzonder Alberto, Reinhard, Björn en Peter voor het dagelijkse gezelschap op de bureau waar het altijd te warm of te koud was. Ook een woord van dank voor de warme ontvangst in de 41 (overigens een prachtig Interbellumgebouw uit de Amsterdamse school) toen we dakloos waren.
Ook mijn vrienden wil ik hier bedanken voor de ontspannen momenten en om me op de meest ongelegen ogenblikken te laten uitleggen wat ik nu weer onderzoek op 'den Unief'.
Dank aan mijn ouders en familie voor de steun doorheen de jaren en in het bijzonder een woordje van dank aan Veerle. Tijdens dit doctoraat groeide je van vriendinnetje tot levenspartner, bedankt voor alles.

Contents

Nederlandse Tekst

Dutch summary - Nederlandse samenvatting

De volgende 40 bladzijden vormen een samenvatting van voorliggend doctoraatswerk. Voor meer details en uitgebreidere achtergrondinformatie verwijzen we naar de Engelse tekst.

1. Inleiding

1.1 Telecommunicatie

1.1.1 Een Beetje Geschiedenis

Door de uitvinding van de elektrische telegraaf in de jaren 1830 beschikte de mensheid voor de eerste keer over een betrouwbare vorm van communicatie die sneller was dan een persoon zelf kon reizen. Vroegere mechanische constructies - zoals de Chappe semafoor in Frankrijk - konden ook hoge snelheden halen (Parijs→Rijsel binnen het uur) maar functioneerden nooit optimaal 's nachts of tijdens een onweer. Ook een oceaan oversteken zou met deze systemen nooit gelukt zijn; wat de telegraaf wel deed toen de overzeese kabel van de *Anglo-American Telegraph Company* in 1866 in gebruik werd genomen.

In de loop der jaren verbeterde de elektrische communicatie en in 1884 verwierf Ernst Siemens een Duits patent op de coaxiale kabel. Het zou meer dan 70 jaar duren vooraleer deze uitvinding in 1956 op de bodem van de oceaan belandde. De eerste transatlantische telefoonkabel, *TAT-1*, was een feit en kon initieel 36 gesprekken tegelijk verwerken.

In de zoektocht naar nieuwe communicatiesystemen, met lager verlies en hogere bandbreedte, was men zich gaan toespitsen op optische systemen. In 1970 fabriceerden onderzoekers van de *Corning Glass Works* de eerste bruikbare glasvezel met een optisch verlies van 98%

Figuur 1: Schema van een optische verbinding met WDM.

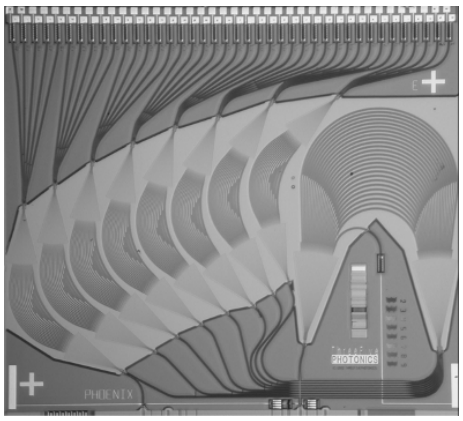
(17 dB) per kilometer. Deze nieuwe technologie bewees zich reeds in 1988 met de verwezenlijking van *TAT-8*, de eerste transatlantische telefoonverbinding gebaseerd op glasvezel, met een initiële capaciteit van 40.000 gesprekken.

Sindsdien heeft men de verliezen van glasvezel verlaagd tot onder $0.15\ dB/km$, wat betekent dat na 20 km nog meer dan de helft van het licht zich in de vezel bevindt. En sindsdien wordt de capaciteit van een telecomverbinding niet langer in een aantal telefoongesprekken uitgedrukt maar in gigabits per seconde (Gbps), omdat het internet en het wereldwijde web voor een spectaculaire stijging van het dataverkeer hebben gezorgd. Telefoongesprekken nemen nog slechts een fractie hier-van voor hun rekening.

1.1.2 Optische Integratie

Om de datastroom door een glasvezel nog te verhogen, verbetert men tegenwoordig vooral de apparatuur aan beide uiteinden. Het sneller aan- en uitschakelen van het licht verhoogt deze datastroom en verkrijgt men door betere lasers, modulatoren en detectoren te ontwerpen. In de laatste jaren komt de grootste verbetering echter door het parallel versturen van verschillende datastromen tegelijk door eenzelfde vezel (WDM), waarbij elke stroom op een andere golflengte wordt gemoduleerd, zie figuur 1. De transatlantische *TAT-14* kabel, in gebruik sinds 2000, bereikt zo een capaciteit van 160 Gbps door 16 kanalen die elk gemoduleerd zijn op 10 Gbps te multiplexen.

De hogere complexiteit van een dergelijk systeem heeft als nadeel dat het aantal onderdelen sterk stijgt. Zo kan een moderne verbinding bestaan uit: multiplexers, demultiplexers, laserdiodes, isolatoren, externe modulatoren, Erbium-gedopeerde versterkers, optische pompen, dispersiecompensatoren, regeneratoren, etc.



Figuur 2: Voorbeeld van een state-of-the-art fotonisch geïntegreerd circuit. Courtesy III-V Photonics BV, an ASIP Inc Company.

Vandaag de dag zit quasi elke component nog in zijn eigen omhulsel met eigen vezelconnector en eigen temperatuurcontroller en worden alle componenten aaneengeschakeld in een lange keten. Het voordeel hiervan is een modulair systeem, waarin een kapotte component snel vervangen is. Het grote nadeel is echter de hoge prijs van een enkele component omdat de kost van het assembleren van de component met de controller en de connector tot 75% van de totale kost kan uitmaken.

Om dit kostprobleem op te lossen, wordt er wereldwijd veel onderzoek verricht naar het combineren van verschillende optische functies in een enkele component. Enkele voorbeelden van deze queeste naar optische integratie zijn het combineren een laser met een isolator, multigolflengtelasers, optische kruisconnecties, het combineren van een versterker met een detector, etc.

Figuur 2 toont een state-of-the-art voorbeeld van een fotonisch geïntegreerd circuit met 9 golflengtesplitsers en 40 fotodetectoren. De onderzoekswereld is ervan overtuigd dat optische integratie de toekomst is en verricht dan ook veel onderzoek op dit gebied. In welke vorm deze integratie uiteindelijk zal geschieden is echter nog geen uitgemaakte zaak.

Een deel van het huidige onderzoek specialiseert zich in monolithische integratie, waarbij alle componenten in eenzelfde materiaalsysteem worden gemaakt, bv. indiumfosfide (InP) of galliumarsenide (GaAs). Actieve componenten (het genereren en detecteren van licht, snelle modulatie) kunnen voorlopig enkel in een halfgeleider met een directe bandkloof worden verwezenlijkt, meestal InP en GaAs. Ac-

tieve en passieve componenten vereisen echter verschillende golfgeleidereigenschappen waardoor complexe technieken als lokale hergroei, quantumlaagvermenging of polarisatiedraaiing nodig zijn voor monolithische integratie. Om de nood aan een directe bandkloof te omzeilen, werken vele groepen aan lichtgeneratie in halfgeleiders met een indirecte bandkloof - meestal het goedkope silicium - door middel van nanostructurering, zeldzame-aarddopering of Ramanversterking.

Monolithische integratie vereist echter compromissen in het ontwerp van een component, waardoor alternatieve integratietechnieken ook onderzocht worden. Passieve componenten kunnen goedkoop gefabriceerd worden in goedgekende materialen als glas en silicium en achteraf gecombineerd worden met actieve componenten uit InP of GaAs. Er bestaan verschillende manieren om ze te combineren, die we zullen aanduiden met hybride en heterogeen.

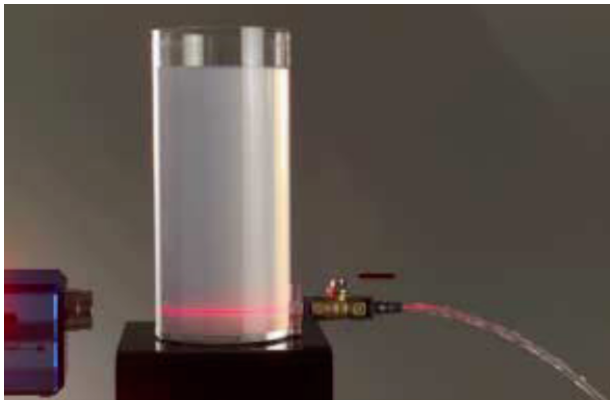
Bij hybride integratie worden de componenten apart gefabriceerd tot een bijna finale vorm en dan gecombineerd. Voorbeelden zijn het samenbrengen van een glazen golflengtesplitser en een rij van optische halfgeleiderversterkers waardoor ze een multi-golflengtelaser vormen. Het assembleren gebeurt meestal handmatig waardoor hybride integratie eerder een onderzoeksmethode vormt dan een bruikbare techniek voor massaproductie.

Heterogene integratie daarentegen kan een techniek voor massaproductie worden maar staat nog in zijn kinderschoenen. Hier worden de verschillende materialen eerst gecombineerd, liefst met een wafer scale techniek. Dan pas worden de componenten verwezenlijkt in het heterogene materiaal. Voorbeelden zijn het plakken van InP films ter grootte van een postzegel op een silicium wafer door gebruik te maken van BCB of de SmartCut techniek.

Welke optische-integratietechniek uiteindelijk tot commerciële componenten zal leiden is nog steeds het onderwerp van onderzoek. Het is wel zeker dat een goedkope technologie voor optische integratie een zeer breed toepassingsdomein zou kennen.

1.2 Lichtgeleiding

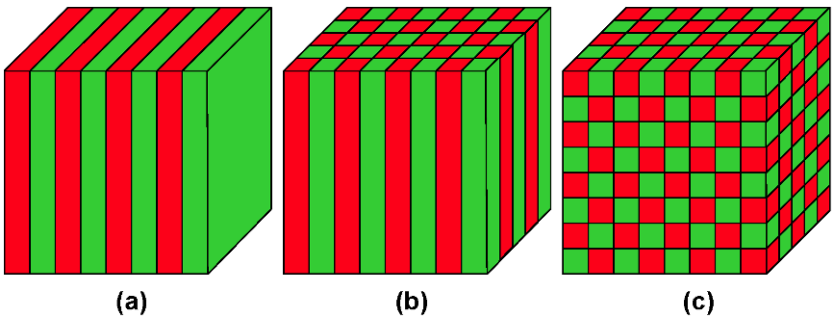
Licht kan door verschillende principes geleid worden. De meest gekende en meest gebruikte manier is om de eigenschap te gebruiken dat licht preferentieel vertoeft in het materiaal met de hoogste brekingsindex, zoals John Tyndall reeds aantoonde in de negentiende eeuw, zie figuur 3.



Figuur 3: Moderne versie van het originele experiment van Tyndall waarin hij aantoonde dat licht door een waterstraal kan geleid worden.

In plaats van water als geleidende laag en lucht als mantellaag, zijn er meer geavanceerde combinaties in gebruik vandaag, zoals gedopeerd / ongedopeerd glas, silicium / glas, silicium / lucht of quaternaire materialen zoals $In_xGa_{1-x}As_yP_{1-y}$ met verschillende (x,y) concentraties. De eigenschap om in het materiaal met de hoogste index te blijven, kan het gemakkelijkst worden verklaard met de wet van Snellius en totale interne reflectie en is daarom ook gekend als geleiding door totale interne reflectie of kortweg TIR-geleiding.

Een andere manier om licht te leiden is door spiegels te gebruiken. Oorspronkelijk heeft men geprobeerd om lichtpijpen te ontwikkelen door de binnenkant van een buis met een metaallaag spiegelend te maken. De (toch) kleine absorptie in het metaal wordt vermenigvuldigd door de talrijke reflecties waardoor de pijpen veel licht lekten. Gelukkig kunnen isolatoren, dit zijn materialen zoals glas die elektriciteit slecht geleiden en daardoor geen licht absorberen, ook gebruikt worden om spiegels te maken. Voor een goede loodrechte reflectie moeten tenminste twee isolatoren periodiek gestapeld worden met een laagdikte van ongeveer $\lambda/4n$ waar λ de golflengte van het licht en n de brekingsindex van de laag is. Aangezien de golflengte van zichtbaar licht kleiner is dan een micrometer, is de fabricage van dergelijke spiegels een hoogtechnologische bezigheid. Een simpele stapeling van lagen als in figuur 4 leidt tot een spiegel voor lichtstralen die onder een welbepaalde hoek of hoekbereik invallen. Om dit bereik te vergroten kan het brekingsindexverschil tussen de lagen verhoogd worden of kunnen de 2 materialen tweedimensionaal gestapeld worden in plaats van



Figuur 4: Spiegels gemaakt uit een **(a)** eendimensionale **(b)** tweedimensionale en **(c)** driedimensionale stapeling van diëlektrische materialen.

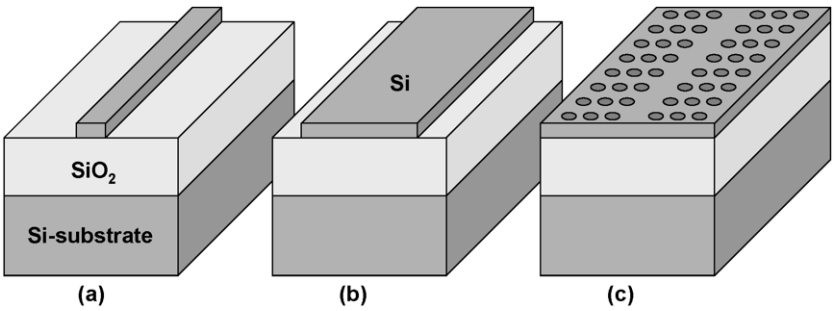
in laagjes, zie figuur 4. Een voor de hand liggende uitbreiding is om een driedimensionale stapeling te maken, zoals in figuur 4 waar een stapeling als die in een zoutkristal gebruikt is.

Dergelijke structuren verplichten fotonen om zich op dezelfde manier te gedragen als een elektron in een atomair kristalrooster en worden daarom fotonische kristallen genoemd. Zoals elektronen in een atomair kristalrooster bepaalde snelheden niet kunnen vertonen voor een bepaalde richting, zo kunnen ook fotonen in fotonische kristallen bepaalde energieën niet verkrijgen voor een bepaalde richting. In analogie met de elektronica waar men dat verboden energiebereik voor elektronen de bandkloof noemt, wordt in fotonica over de fotonische bandkloof (FBK) gesproken. In tegenstelling tot metaalspiegels, kan men deze FBK-spiegels wel succesvol gebruiken voor lichtgeleiding, die we dan eveneens FBK-geleiding zullen noemen.

Alhoewel in theorie zowel een eendimensionale als een driedimensionale stapeling rond een golfgeleider op een optische chip kunnen gewikkeld worden, gebeurt dit in de praktijk bijna enkel met tweedimensionale spiegels zoals in figuur 5(c), omdat een dergelijke structuur compatibel is met de standaard fabricage.

1.3 Optische Overgangen en Koppeling

In fotonische chips kunnen verscheidene golfgeleidertypes, die onderling verschillen van grootte en van geleidingsprincipe toch samen bestaan. Een mogelijk voorbeeld rond het geleidingsprincipe is het gebruik van een fotonische-kristalgolfgeleider om een kleine verliesloze bocht te maken en een TIR-golfgeleider voor het verbinden van ver-



Figuur 5: Verscheidene types golfgeleider die samen op een fotonische chip kunnen voorkomen. (a) smalle TIR golfgeleider (b) brede TIR golfgeleider (c) FBK golfgeleider.

scheidene bochten. Voor het samengaan van golfgeleiders met een verschillende vorm denkt men aan een monomodale golfgeleider voor het bochtwerk en iets bredere multimodale golfgeleiders, met minder verlies, voor het rechte transport van licht in een fotonische chip.

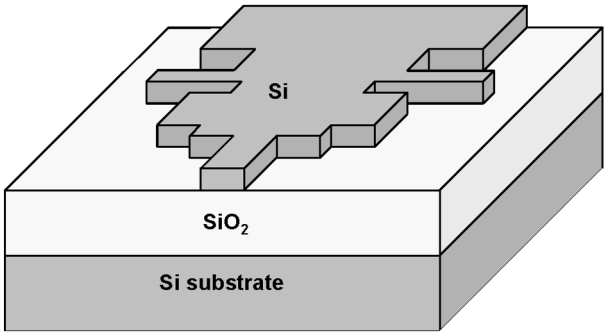
Maar ook in de verbinding van een optische chip met de buitenwereld moeten lichtbundels van verschillende vorm en grootte met elkaar gekoppeld worden, zoals een glasvezel met een vlakke golfgeleider of een halfgeleider-laserdiode met een vlakke golfgeleider.

Omdat het licht geleid moet worden tussen al deze verschillende golfgeleiders is het belangrijk dat de verbindingen ontworpen worden om zo weinig mogelijk verlies te veroorzaken. Een optische component die een dergelijke verbinding tot stand brengt, noemen we een lichtbundelomvormer of, hierna, kortweg een omvormer.

1.4 Doel en Overzicht van het Werk

In dit werk leggen we ons toe op de studie van korte vlakke koppelingen tussen golfgeleiders. Heden gebruikt men graduele (lineaire of parabolische) overgangen tussen smalle en brede golfgeleiders als in figuur 5. We zullen echter van nul af aan beginnen en een optimalisatie verrichten op de structuur in figuur 6. Zo zal er bepaald worden welke dimensies de stukken golfgeleider, die tussen de smalle en de brede golfgeleiders geplaatst zijn, moeten vertonen om tot een goede koppeling te leiden.

Na het onderzoek hoe een performante omvormer moet ontworpen worden, zullen dezelfde algoritmes toegepast worden om de koppeling



Figuur 6: Een generieke koppelstructuur bestaande uit verscheidene golfgeleidersecties tussen een smalle en een brede golfgeleider.

tussen golfgeleiders met een verschillende geleidingsprincipe, meer bepaald tussen een TIR en een FBK golfgeleider.

De beslissing om deze studie toe te spitsen op vlakke omvormers mag niet als een beperking gezien worden. Verderop wordt er uitgelegd hoe een vlakke omvormer samen met een verticale vezelkoppeling tot een compacte driedimensionale omvormer kan leiden.

De tekst in dit boek is op de volgende wijze geordend. In deel 2 worden de geschiedenis en bestaande literatuur van lichtbundelomvormers kort besproken waarbij er speciale aandacht is voor het recentere werk waarvan deze studie een logisch voortvloeisel is.

Deel 3 geeft een kort overzicht van de optische-simulatietechnieken en de wiskundige optimalisaties die verderop gebruikt zijn. De eigenlijke optimalisatie en de resulterende ontwerpen voor omvormers worden uitgebreid besproken in deel 4 en vergeleken met de bestaande omvormers. Drie verschillende methodes om een omvormers te construeren worden getest en vergeleken met elkaar.

Een korte inleiding in de realisatie van de ontwerpen in silicium-op-isolator wordt gegeven in deel 5. De meetmethode wordt kort besproken waarna de beste metingen en de gebruikte dataverwerking volgen.

In het laatste deel worden de besluiten van dit werk samengebracht en enkele voorstellen voor het vervolg op deze studie gedaan.

1.5 Publicaties

De resultaten van dit werk werden gepubliceerd in verscheidene papers en voorgesteld op enkele conferenties. Voor de publicaties zelf verwijzen we graag naar de Engelse tekst. Ze staan ook online vermeld

in de sectie publicaties op de website van de onderzoeksgroep fotonica: http://photonics.intec.ugent.be.

2. Lichtbundelomvormers

Het onderzoek naar golfgeleiders krijgt een stevige duw in de rug door een publicatie van Kao en Hockham in 1966, waarin ze uiteenzetten welke grootte en brekingsindex de kern en de mantel van een glasvezel moeten hebben en waarin ze aantonen dat een optisch verlies van minder dan $20\ dB/km$ haalbaar moet zijn, door de onzuiverheden uit het glas te halen. Deze paper overtuigt veel onderzoeksgroepen om te starten in dit beloftevolle onderzoeksgebied, dat door de recente uitvinding van de laser eindelijk zou kunnen leiden naar de heilige graal van de telecommunicatie: een lange-afstandsverbinding met weinig verlies en een grote datastroom. In Bell Labs schrijft Dietrich Marcuse vanaf 1969 een aantal theoretische verhandelingen rond het effect van een niet-vlakke grenslijn tussen de kern en de mantel.

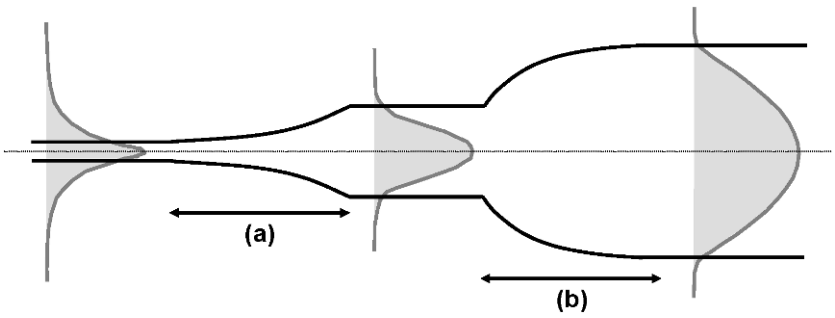
2.1 Literatuur van Algemene Optische Omvormers

Een wigvormige omvormer wordt het eerst vermeld door Marcuse, die afleidt wat het stralingsverlies is van een kleine breedteverandering in een golfgeleider en die kennis gebruikt om het stralingsverlies bij een grote geleidelijke verandering uit te drukken. Hij bespreekt het geval waar een reeds monomodale golfgeleider nog verder versmalt en besluit dat om het stralingsverlies zoveel mogelijk te verspreiden over de gehele omvormer het versmallen sneller mag gebeuren in het brede gedeelte. Deze omvormer noemen we een omgekeerde omvormer.

Een tweede belangrijke paper is de publicatie van Winn en Harris in 1975 waarin ze de koppeling van een monomodale naar een multimodale golfgeleider beschrijven aan de hand van hoornvormige omvormers en met behulp van gekoppelde-modetheorie. Ook hier is de conclusie dat de koppelefficiëntie kan verhoogd worden door de vorm van de omvormer goed te kiezen en ze stellen de volgende vorm voor:

$$w(z) = w_0 + 2z \cdot \tan(\theta/2) + A\left(1 - \exp\left(-z/L\right)\right) \tag{1}$$

met w_0 de breedte van de golfgeleider aan de ingang, L de lengte en θ de volledige hoek van de omvormer. Deze eerste twee termen rechts van de gelijkheid beschrijven een lineaire omvormer, de derde term



Figuur 7: Verschillende omvormers: **(a)** een omgekeerde voor koppeling tussen monomodale golfgeleiders met slechte en goede insluiting en **(b)** een reguliere omvormer voor koppeling tussen een mono- en een multimodale golfgeleider

verbreedt deze vorm vooral aan de smalste kant. Het wordt ook numeriek aangetoond dat deze vorm effectief voor een betere koppelefficiëntie zorgt. Dit soort omvormer noemen we een reguliere taper, zie figuur 7.

Als een alternatief voor de bovenstaande vergelijking stellen Burns *et al.* in 1977 als eerste de parabolische hoornvorm voor met als vergelijking:

$$w(z) = \sqrt{\frac{2\alpha\lambda_0}{n_{core}}z + w_0^2} \tag{2}$$

met α een constante die kleiner is of gelijk aan 1, en soms de adiabatische parameter wordt genoemd. Aangezien deze vergelijking wordt afgeleid aan de hand van een stralenmodel, verfijnen de auteurs ze verder tot de uitdrukking:

$$w(z) = \sqrt{2\alpha\lambda_g z + w_0^2} \tag{3}$$

met λ_g de echte golflengte van de lokale grondmode. Deze keer gebruiken de auteurs eigenmode-expansie en gekoppelde-modetheorie. Om rigoreus een parabool te verkrijgen moet λ_g plaatsonafhankelijk zijn, waardoor je moet veronderstellen dat zowel de grond als de eerste mode ver weg zijn van cut-off. Hierdoor wordt de bovenstaande vergelijking enkel geldig voor koppeling tussen brede multimodale golfgeleiders met een sterke insluiting. De auteurs vermelden verder dat in het algemene geval de optimale hoornvorm niet strikt parabolisch zal zijn.

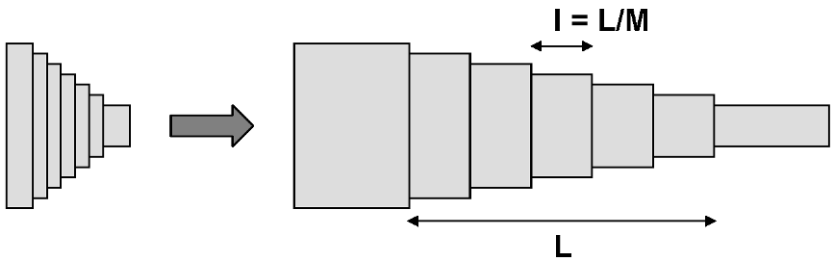
Bij ons weten is de paper van Milton *et al.* uit 1977 de eerste die het woord adiabatisch gebruikt in de context van optische bundelomvormers. Afkomstig uit de thermodynamica, werd het later in de kwantummechanica gebruikt om een systeem te beschrijven dat zodanig traag varieert dat de energie de verandering niet voelt en in dezelfde toestand blijft, ook tijdens de toestandsverandering. In golfgeleidertheorie, vertaalt zich dit naar de situatie waar alle energie in de grondmode blijft, ook als golfgeleider zelf verandert. We zullen het woord adiabatisch enkel in deze betekenis gebruiken en nooit in de ietwat ruimere betekenis van een systeem waar geen energie verloren gaat ongeacht door welke toestanden deze energie gaat.

2.2 Driedimensionale Bundelomvormers

De eerste papers over omvormers gebruikten meestal vlakke golfgeleiders voor de berekeningen. Toch werden de eerste experimenten verricht op driedimensionale omvormers aangezien vooral de koppeling tussen bijvoorbeeld een laser en een glasvezel toen onderzocht werd. Een golfgeleider in beide transversale richtingen van grootte laten veranderen is technologisch echter niet eenvoudig, waardoor men op zoek gegaan is naar methodes om een golfgeleidermode in beide richtingen te laten veranderen terwijl de golfgeleider zelf dat enkel in de horizontale richting doet, en de fabricage dus compatibel blijft met de gebruikelijke technologie. Voor een overzicht van dit onderzoek verwijzen we naar het overeenkomstige hoofdstuk in de Engelse tekst rond geïntegreerde bundelomvormers en mode-expansie met vlakke technieken. Het is in deze context dat een eerste keer een puur wiskundige optimalisatie verricht werd op een omvormer en het is op dit ingeslagen pad dat we willen voortgaan. In de Engelse tekst wordt eveneens een klein overzicht gegeven van het gebruik tot nu toe van optimalisatie in geïntegreerde optica.

2.3 Tweedimensionale Omvormers

Deze paragraaf handelt over omvormers waar de golfgeleidermode in de verticale richting quasi onveranderd blijft bij doorgang doorheen de omvormer.



Figuur 8: Uitrekken tot een lange getrapte omvormer.

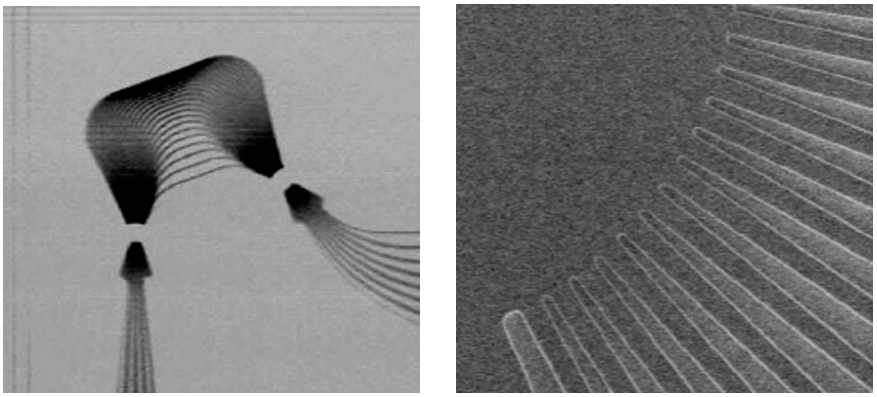
2.3.1 Literatuuroverzicht

Sinds de eerste papers van paragraaf 2.1 is er weinig specifiek onderzoek verricht rond tweedimensionale omvormers omdat het grootste knelpunt vooral in de koppeling zat tussen een component en de buitenwereld. Het aantal koppelingen binnenin een optische component was zeer klein waardoor de grootte ervan er niet echt toe deed en waardoor lange lineaire omvormers gebruikt konden worden. De enkele studies die toch verricht werden worden hieronder besproken.

Ter vervanging van een gladde parabool heeft Kiseliov de beste manier bestudeerd om een parabool te benaderen een stuksgewijze lineaire vorm. Dit werk was oorspronkelijk bedoeld voor millimetergolven maar kan volledig overgeplaatst worden naar optica. Gegeven een aantal lineaire secties, worden de coördinaten van de eindpunten van deze secties berekend. Opnieuw is deze theorie enkel toepasbaar op golfgeleiders die multimodaal zijn met een sterke insluiting.

Op lithografische maskers, wordt een gladde parabool meestal benaderd door een getrapte parabool met weliswaar kleine trapjes. Afhankelijk van de resulotie van de lithografie, kunnen deze trapjes al dan niet in de echte component nog zichtbaar zijn. Felici *et al.* hebben naar het mogelijke effect van getrapte omvormers op de uiteindelijke transmissie gekeken. Het is allang geweten dat voor omvormers rond de adiabatische limiet deze trapjes vrij groot mogen zijn voor numerieke berekeningen. Als deze grove discretisatie echter wordt toegepast op zeer lange omvormers, als in figuur 8, kunnen er bijkomende verliezen optreden en kan in extreme gevallen de transmissie zelfs naar nul vallen.

In een andere publicatie van Felici *et al.* wordt een omvormer als invers probleem bestudeerd. In plaats van een bepaalde vorm voorop te stellen en dan te bestuderen hoe die functioneert, werken de au-



(a) Golflengtesplitser in SOI

(b) Plaatgolfgeleider en de fasegolfgeleiders

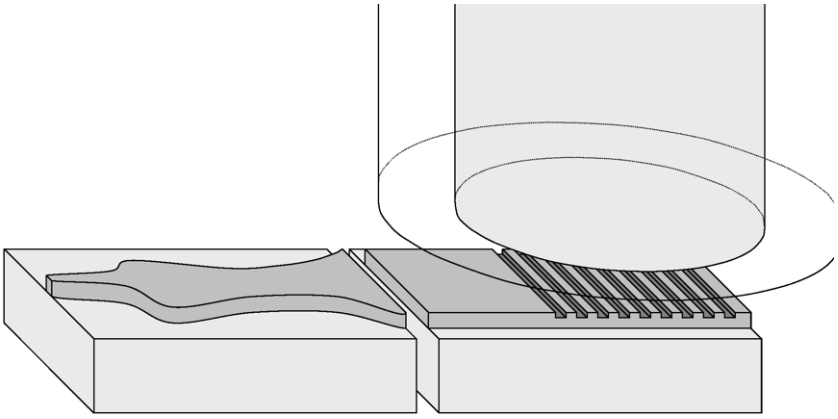
Figuur 9: SEM-beelden van een golflengtesplitser (een AWG) in SOI en een close-up van de overgang tussen de plaatgolfgeleider en de fasegolfgeleiders.

teurs andersom en proberen ze een bepaalde brekingsindexverdeling te vinden die de grondmode zo goed mogelijk tussen twee verschillende golfgeleiders koppelt. Er worden wel bepaalde randvoorwaarden opgelegd aan deze verdeling, onder andere dat ze realiseerbaar is met bekende materialen. Dit werk is sterk mathematisch georiënteerd en heeft als belangrijkste conclusie dat er veel meer mogelijke brekingsindexverdelingen bestaan, dan bijhorende veldverdelingen (op een mathematische wijze zoals $\infty^2 > \infty$). Hierdoor is het nodig om genoeg voorwaarden aan de indexverdeling op te leggen, anders is convergentie tijdens de optimalisatie helemaal niet gegarandeerd.

2.3.2 Toepassingen van Tweedimensionale Omvormers

Vandaag de dag worden omvormers vooral gebruikt om een enkele optische component (zoals een laser) naar een glasvezel te koppelen. Geïntegreerde optica zal dit plaatje echter veranderen in de toekomst. Een verbinding naar de buitenwereld zal nog steeds nodig zijn maar door de integratie van verschillende optische functies in een chip, worden twee-dimensionale omvormers ook essentiële componenten. Dit is het geval van zodra er twee verschillende golfgeleiders zo verliesloos mogelijk aan elkaar moeten passen.

Een mogelijke toepassing is in een complex optisch circuit dat bijvoorbeeld draadgolfgeleiders (ongeveer 0.5 µm breed) gebruikt. Om



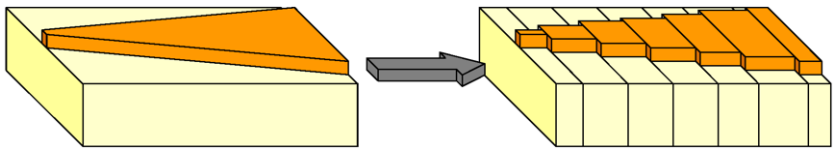
Figuur 10: Verticale roosterkoppeling tussen een brede golfgeleider en een glasvezel. Een conceptuele compacte omvormer is ook zichtbaar in de figuur.

een kleine bocht te kunnen maken moet de draad strikt monomodaal zijn, wat wil zeggen dat er een enkele geleide mode in kan bestaan, en een sterke insluiting vertonen. In SOI komt dit volgens Vlasov overeen met een breedte van ongeveer 445 nm, waarbij de propagatieverliezen vrij hoog zijn. Lange rechte stukken golfgeleider, zoals wachtlijnen of verbindingen tussen verscheidene functies, zijn gebaat met een kleiner verlies en moeten daardoor breder gemaakt worden. Een omvormer is dan nodig om de smalle en bredere golfgeleiders aan elkaar te zetten.

Een andere toepassing van omvormers in een optische chip kan binnenin een bepaald soort golflengtesplitser, een AWG, zijn, zie figuur 9(a). De overgang tussen de plaatgolfgeleider en de fasegolfgeleiders moet met zorg ontworpen worden om reflectie te onderdrukken en zoveel mogelijk licht in de fasegolfgeleiders te krijgen. Een recente techniek om deze overgang te verbeteren is om twee verschillende etsdieptes te gebruiken. We denken dat een compacte tweedimensionale omvormer ook in deze situatie zijn nut kan hebben.

Ook aan de overgang tussen conventionele TIR-golfgeleiders en fotonische kristal of FBK-golfgeleiders kunnen tweedimensionale omvormers gebruikt worden. Voor een overzicht van het recent onderzoek in dit gebied verwijzen we graag naar de Engelse teksst.

De recente succesvolle demonstratie van een breedbandige verticale vezelkoppeling door Taillaert *et al.*, zie figuur 10, en de mogelijkheid om deze component in te bouwen in elk materiaalsysteem met een potentieel groot brekingsindexverschil, heeft ook een extra toepassing geschapen voor tweedimensionale omvormers. Hier is er ook een



Figuur 11: Discretisatie van een structuur in lagen die invariant zijn in de z-richting.

groot verschil in breedte tussen de brede golfgeleider waarin het rooster gedefinieerd wordt en de smalle golfgeleider (12 µm tegenover 0.5 µm) die de start vormt voor de rest van het circuit. Door dit grote verschil zal een conventionele omvormer al snel veel plaats innemen en kan een compactere versie een echte plaatsbesparing vormen.

3. Methodes

3.1 Eigenmode-Expansiemethode

Alle simulaties die verderop in dit werk beschreven worden, zijn uitgevoerd met een eigenmode-expansie methode of EEM.

Om de vergelijkingen van Maxwell voor een bepaalde structuur op te lossen met eigenmode-expansie, moet er eerst een propagatierichting gekozen worden. Aangezien we tweedimensionale omvormers bestuderen, valt deze propagatierichting vanzelfsprekend samen met de optische as van de beide golfgeleiders, en zullen we dit ook altijd als de z-as aanduiden. Het simulatiegebied moet ook eindig zijn, wat we in orde brengen door het in een rechthoekige box te plaatsen met magnetische of elektrische wanden. Dit zijn artificiële wanden, die sommige componenten van het elektrisch of magnetisch veld tot nul dwingen. Hierdoor is er ook perfecte reflectie aan deze wanden, zodat speciale randvoorwaarden worden ingevoerd om deze reflecties kwijt te geraken. De meest gebruikte absorberende randvoorwaarde in een hele reeks elektromagnetische simulaties is vandaag de perfecte aangepaste laag of PML. De combinatie van elektrische of magnetische wanden en PML biedt het voordeel van een eindig simulatievolume terwijl van binnenuit gezien dit volume quasi ononderscheidbaar is van een open volume.

Vervolgens moet de structuur in lagen verdeeld worden die invariant zijn in de z-richting, zoals in figuur 11. Dit lijkt op het eerste zicht misschien een beperking van de methode maar door de discretisatie-

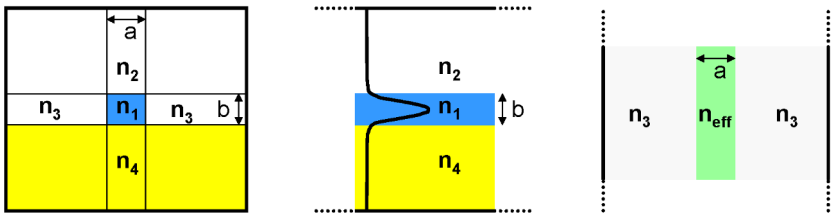
stap voldoende klein te kiezen kan elke vloeiende structuur toch zo dicht als wenselijk benaderd worden. De transversale brekingsindexverdeling in een laag wordt ingevoerd in de vectoriële Helmholtz vergelijking en de eerste N (complexe) eigenwaarden worden hiervan bepaald. Deze eigenwaarden stellen we voor met β_i en worden ook de propagatieconstanten genoemd. Door de relaties voor elektromagnetische velden aan een interface te gebruiken, berekenen we bij elke propagatieconstante een bijhorende eigenmode, of modeprofiel. Deze eigenmodes of kortweg modes zijn de natuurlijke manieren waarop licht kan bewegen in een golfgeleider. Als alle energie in een welbepaalde mode zit in een perfecte golfgeleider, blijft ze ook in die mode.

Voor elke overgang tussen twee secties kan er een verstrooiingsmatrix opgesteld worden die de verbanden bevat tussen de modes die in beide richtingen propageren aan de ene kant en de andere kant van de overgang. Ook de propagatie van de modes binnenin een sectie kan door een matrix beschreven worden. Voor alle achtereenvolgende secties en overgangen ertussen, kunnen deze matrices dan op de juiste manier gecombineerd worden tot een verstrooiingsmatrix die de gehele structuur beschrijft, een meer gedetailleerde uitwerking van dit proces staat beschreven in de Engelse tekst. Een van de elementen van de totale verstrooiingsmatrix drukt uit hoe de grondmode aan het begin van de structuur uiteindelijk zal koppelen naar de grondmode aan de uitgang van de structuur. Het is de waarde van dit element, of eerder de modulus van het kwadraat ervan, die ons onmiddellijk de koppelefficiëntie van de structuur zal geven. Om de nauwkeurigheid van een bereking d.m.v. eigenmode-expansie te verhogen, wordt het aantal gebruikte modes verhoogd tot een zeker convergentiepunt.

3.2 Effectieve-Indexmethode

Driedimensionale berekeningen vereisen veel meer tijd dan tweedimensionale, waardoor ze vandaag nog onbruikbaar zijn voor de optimalisatie door middel van vele berekeningen van een structuur. Hierdoor worden er meestal tweedimensionale berekeningen uitgevoerd om nieuwe principes en componenten te demonstreren.

Om een echte structuur in een equivalente 2-D structuur om te zetten, zou je een oneindig dun plakje doorheen het midden van de structuur kunnen bedenken, evenwijdig aan het vlak van de component. Het blijkt echter dat dit niet de beste keuze is en dat de effectieveindexmethode een betere optie is. Deze methode werd het eerst be-



Figuur 12: Illustratie van de effectieve-indexmethode. **Links:** originele doorsnede loodrecht op de propagatie-as. **Midden:** middelste verticale lagen om de grondmode te berekenen. **Rechts:** uiteindelijke effectieve-indexverdeling. Een stippellijn stelt een oneindige voortzetting voor.

schreven voor rechthoekige optische golfgeleiders door Hocker en Burns in 1977. In plaats van de brekingsindex te gebruiken van een simpele horizontale snede door de structuur, verdelen ze de structuur in verticale secties, loodrecht op de voortplantingsrichting, en in elke sectie definiëren ze een gemiddelde brekingsindex, als volgt. De doorsnede loodrecht op de propagatie-as wordt verdeeld in verticale secties waarbinnen de brekingsindex horizontaal niet verandert, als in figuur 12. De sectie die de kern van de golfgeleider bevat wordt getransformeerd in een plaatgolfgeleider door alle secties oneindig lang uit te rekken. Vervolgens wordt de grondmode berekend, wat in een propagatieconstante resulteert. Via de formule $\beta = 2\pi n_{eff}/\lambda_0$ hoort er bij deze β een effectieve index, die je kan beschouwen als de gemiddelde brekingsindex die het licht tegenkomt. In een volgende stap wordt deze effectieve index toegekend aan de gehele verticale laag, alsof het een fysische brekingsindex was. Het is noodzakelijk om tijdens deze procedure de veldoriëntatie in de originele golfgeleider in de gaten te houden en telkens de juiste randvoorwaarden te gebruiken. Als de andere verticale secties ook tot plaatgolfgeleiders kunnen gerekt worden, dan wordt de procedure herhaald in elke sectie, zoniet dan wordt de brekingsindex (n_3) naast de kern gebruikt, zie opnieuw figuur 12.

3.3 Genetische Algoritmes

Genetische algoritmes werden voor het eerst onderzocht door John Holland ergens in het midden van de jaren 60 wanneer hij probeerde om artificiële intelligentie te creëren door het verloop van natuurlijke evolutie te gaan nabootsen aan de hand van computers. Meer bepaald liet hij computerprogrammaatjes natuurlijk evolueren om complexe problemen op te lossen, zoals het ordenen van een rij getallen. IN de loop der

jaren heeft de methode zijn weg gevonden vanuit de AI-gemeenschap naar wiskunde en de toegepaste wetenschappen, waar ze nu gebruikt wordt als optimalisatietechniek.

In pseudo-code ziet een genetisch algoritme er als volgt uit, met gebruik van de woordenschat schatplichtig aan de biologische evolutie:

Kies een beginpopulatie
Herhaal
    Evalueer de aangepastheid van elk individu
    Selecteer de best geplaatste individuen voor reproductie
    Laat willekeurige koppels paren
    Pas de crossover-operator toe
    Pas de mutatie-operator toe
    Voeg de nieuwe individuen toe aan de generatie
    Verwijder de slechtst aangepaste individuen uit de generatie
Tot een zekere stopconditie

Individu: een bepaalde set van parameterwaarden $(x_1, x_2, \cdots, x_n)$ die een structuur volledig beschrijft, in ons geval een omvormer. In het zuiverste genetische algoritme worden niet de parameters zelf maar een binaire representatie ervan gebruikt, net zolas in de biologie niet de eiwitten zelf van ouder naar kind worden doorgegeven maar de DNA encodering van deze eiwitten. We zullen echter werken met de fysische parameters zelf (zoals de breedte van een golfgeleider) aangezien dit meer controlemogelijkheden biedt gedurende de opbouw en het gebruik van het algoritme.

Populatie: een groep individuen. De beginpopulatie wordt meestal willekeurig samengesteld met weinig beperkingen. Deze 'domme' creatie van individuen is belangrijk omdat ze toelaat om snel, alhoewel grofweg, de gehele parameterruimte af te scannen.

Aangepastheid (Eng. fitness): de bekwaamheid van een individu. In natuurlijke evolutie zal het individu overleven dat het best aangepast is aan zijn omgeving. In een genetisch algoritme is de fitness-functie, de functie die geoptimaliseerd wordt. In ons geval zal de transmissie van de grondmode van een omvormer, de fitness van die omvormer zijn. Waar mogelijk wordt het fitness-interval gekozen als $[0 \ldots 1]$, zo kunnen gemakkelijk hogere machten van de fitness gebruikt worden

om de beste individuen nog meer te onderscheiden van elkaar.

Selectie: de manier waarop individuen koppels vormen alvorens te reproduceren. In een genetisch algoritme gebeurt dit op een willekeurige wijze maar meestal hebben de beste individuen een grotere kans om verkozen te worden. In dit werk hebben we roulettewielselectie gebruikt als kiesalgoritme.

Reproductie: vanuit twee individuen worden er een paar nieuwe gegenereerd. Reproductie is het gevolg van crossover en mutatie samen.

Paren: zie Reproductie

Crossover of uitwisseling: de operator die twee individuen informatie laat uitwisselen om nieuwe individuen te creëren, in ons geval de vermenging van de twee parameter sets

$$(x_1, x_2, \cdots, x_n) \,\&\, (y_1, y_2, \ldots, y_n) \tag{4}$$

om twee nieuwe te genereren

$$(x_1, x_2, \ldots, y_{n-1}, y_n) \,\&\, (y_1, y_2, \ldots, x_{,n-1}, x_n). \tag{5}$$

Er bestaan verscheidene crossover operatoren, zoals de uniforme, en de een- en tweepunts crossover. De eerste twee werden gebruikt in dit werk en worden gedetailleerd uiteengezet in de Engelse tekst.

Mutatie: na crossover worden de twee nieuwe parametersets lichtjes gewijzigd, onafhankelijk van elkaar. Dit is de tegenhanger van accidentele DNA-mutaties door kopieerfouten of externe straling. Mutatie zorgt ervoor dat het genetisch algoritme niet vastloopt in lokale minima van zodra de diversiteit binnen een populatie begint af te nemen. De uiteindelijke individuen, na crossover, zien er uit als volgt: $(x_1 + \Delta_1, x_2, \ldots, y_{n-1} + \Delta_2, y_n)$ en $(y_1, y_2 + \Delta_3, \ldots, x_{,n-1}, x_n)$ na de mutatie. Mogelijke mutatie-operatoren zijn de bitflip, grensmutatie, uniforme, gaussiaanse, etc. Aangezien we de gaussiaanse gebruikte hebben wordt die ook gedetailleerd besproken in de Engelse tekst.

Generatie: alle individuen die leven op een bepaald moment tijdens het algoritme. In tegenstelling tot bij natuurlijke evolutie, laten we toe dat ouders en hun nageslacht tot dezelfde generatie behoren.

Verwijdering: We werken met een vaste generatiegrootte, waardoor sommige individuen moeten verwijderd worden op het einde van elke iteratie. Dit afsterven gebeurt door de individuen die het slechtst aangepast zijn te verwijderen.

Stopconditie: We willen dat de optimalisatie gestopt wordt binnen een eindige tijd, waardoor we het zelf kunstmatig moeten stilzetten. Dit kan gebeuren na een vast aantal iteraties of wanneer de variatie (in fitness) binnenin een generatie kleiner wordt dan een vooraf ingestelde waarde. Andere stopcondities zijn natuurlijk denkbaar.

4. Simulatieresultaten

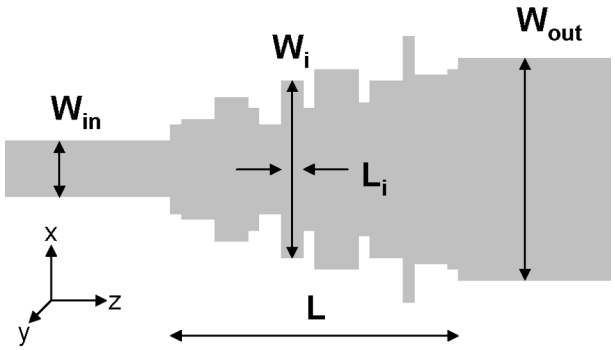
In plaats van te bestuderen welke vorm een omvormer moet hebben om zo weinig mogelijk licht te koppelen naar hogere-ordemodes en vanuit deze studie dan te gaan kijken hoelang een omvormer moet zijn om bijna verliesloos te werken, zullen we hier de omgekeerde weg volgen. We vertrekken van een omvormer met een vaste lengte en gaan de vorm manipuleren totdat het verlies zeer klein geworden is. Op deze wijze hopen we tot ontwerpen te komen die bruikbaar zijn in de praktijk en toch veel kleiner dan de bestaande ontwerpen.

4.1 Inleiding

Aangezien we eigenmode-expansie toepassen om omvormerstructuren te berekenen, wat impliceert dat we die in z-invariante secties moeten verdelen, komt de constructie van een efficiënte omvormer overeen met het bepalen van goede parameters voor deze golfgeleidersecties. Figuur 13 toont een voorbeeld van een omvormer in zijn meest algemene vorm. Tussen de ingangsgolfgeleider met breedte W_{in} en de uitgangsgolfgeleider met breedte W_{out} worden er N golfgeleidersecties geplaatst met een breedte W_i en een lengte L_i zodat

$$\sum_{i=1}^{N} L_i = L \tag{6}$$

waar L de totale lengte van de omvormer is. Voor rigoreuze 3-D berekeningen moet de verticale lagenstructuur volledige in rekening gebracht worden. In 2-D simulaties, na een effectieve-indextransformatie, ziet



Figuur 13: Structuur van een omvormer in zijn meest algemene vorm.

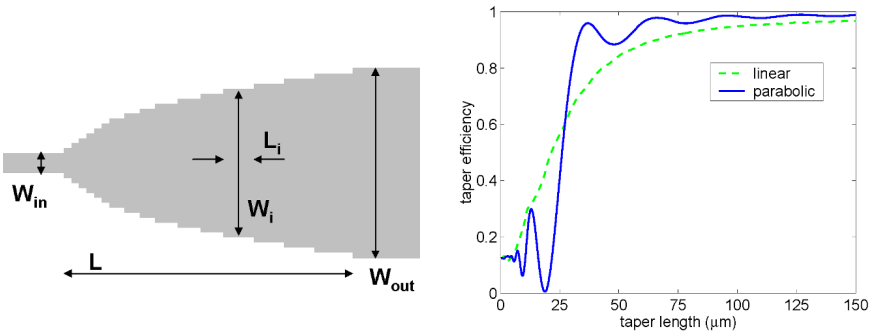
de simulatie eruit als in figuur 13, waar slechts twee brekingsindices gebruikt worden, een voor de kern en een voor de mantel.

De optimalisatie van de structuur zal op verschillende manieren aangepakt worden. De initiële voorwaarden kunnen verschillen. Soms wordt er gestart van een parabool, die dan verbeterd wordt. Soms wordt er vertrokken van compleet willekeurige waarden. Ook de parameters die tijdens de optimalisatie worden veranderd, kunnen varieren. Zo worden de 2 parameters (W_i, L_i) die een sectie beschrijven eens direct geoptimaliseerd, ergens anders wordt de continue omhullende van deze secties veranderd. Ook in het genetische algoritme zijn er soms kleine variaties, zoals de expliciete vorm van de crossover of de mutatie-operatoren.

In dit werk wordt telkens dezelfde polarisatie gebruikt, namelijk die waarbij het dominante transversale elektrische veld in de x-richting ligt, d.i. de breedterichting in figuur 13. Deze polarisatie wordt in de literatuur meestal TE genoemd.

4.2 Gewijzigde Gediscretiseerde Parabool

We vertrekken hier van een gekende en werkende structuur: de parabolische omvormer. De gebruikte breedte van de ingangsgolfgeleider is 0.56 µm en die van de uitgangsgolfgeleider 10.0 µm. Voor de eigenmode-expansieberekening moet de continue vorm gediscretiseerd worden in z-invariante secties zoals in figuur 14(a). Als het aantal secties groot genoeg gekozen wordt, geeft de berekening hetzelfde resultaat als de oorspronkelijke gladde omvormer. In ons geval waren 200 stappen genoeg om geen residueel effect van de discretisatie meer te



(a) Gediscretiseerde parabool (b) Efficiëntie voor lineair en parabolisch

Figuur 14: Links: parabolische omvormer gediscretiseerd op een vaste breedte (x) stap. **Rechts:** transmissie-efficiëntie voor een lineaire en parabolische omvormer, 200 golfgeleidersecties, W_{in}=0.56 µm, W_{uit}=10.0 µm, λ=1.55 µm.

merken. De efficiëntie van een omvormer is berekend als functie van de lengte. Voor zeer korte lengtes is de waarde identiek aan de efficiëntie als je de twee golfgeleiders gewoon aan elkaar zou zetten. Voor zeer grote lengtes convergeert de efficiëntie naar 1.0 of 100%. Tussenin zien we in figuur 14(b) een bijna monotone stijgende curve voor het lineair geval en een periodieke en veel onregelmatiger curve voor de parabool. Zoals theoretisch bepaald, wordt de parabool reeds adiabatisch voor kortere lengtes dan de lineaire omvormer. De gemeten periodiciteit (rond 30 µm) in de efficiëntie voor grotere lengtes, komt goed overeen met de (dubbele) koppellengte tussen de grondmode en de volgende even mode:

$$L_{coupling} = \frac{2\pi}{\bar{\beta}_0 - \bar{\beta}_2} \tag{7}$$

met

$$\bar{\beta}_i = \frac{\int_0^L \beta_i dL}{L} \tag{8}$$

waarbij β_i de propagatieconstante van de i^{de} mode en L de totale lengte van de omvormer is. De eerste twee grote efficiëntiepieken (37 en 66 µm) worden gebruikt als vertrekpunten in de optimalisatie, en dit gediscretiseerd met een vaste x-stap en in totaal 60 secties.

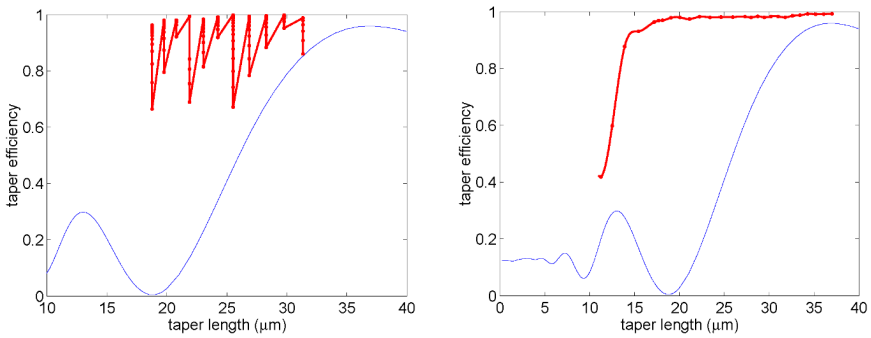
We gebruiken hier een speciaal optimalisatie-algoritme. Men vertrekt bij de eerste sectie aan de linkerkant (W_1, L_1) en berekent de transmissie van de omvormer voor $W_1' = W_1 \pm dx$. Als een van de veranderingen beterschap brengt in de transmissie, verandert W_1 naar deze

waarde. Dan wordt er gekeken met W_1 een dx-stap verder om te zien of er nog verdere verbetering mogelijk is, totdat W_1 zich in een lokaal maximum bevindt. Hetzelfde proces wordt herhaald voor W_2, W_3, ... totdat de rechterkant van de omvormer bereikt is. Door de veranderingen aan sommige andere breedtes, is het niet meer gegarandeerd dat W_1 in het lokale maximum zit, waardoor het scannen met dx-stapjes herbegint en nu zelfs kan leiden tot veranderingen in de tegengestelde breedterichting. Bemerk echter dat met elke verandering aan de omvormer de transmissie verbetert. Na W_1, ondergaan alle andere breedtes nog eens dezelfde procedure. Het aflopen van alle secties van links naar rechts gaat door tot een zekere stopconditie. Dit kan een maximaal aantal iteraties, een drempeltransmissie of een te kleine verbetering in de laatste iteratie zijn. Tot nu toe bleven de lengtes L_i van de golfgeleidersecties ongewijzigd. Na de laatste links-rechtsbeweging worden alle lengtes in een klap gewijzigd aan de hand van de relatie:

$$L_i' = a_s \cdot L_i \tag{9}$$

waar a_s een verkortingsfactor is, een beetje kleiner dan een. Na de lengteverandering, herbegint de breedte-optimalisatie opnieuw. Dit gehele proces gaat door tot de omvormer de gewenste lengte bereikt heeft.

De gehele optimalisatie is gekenmerkt door de initiële lengte, de dx-stap, de verkortingsfactor a_s en de tussentijdse stopvoorwaarde. Deze getallen kunnen niet zomaar gekozen worden en hun exacte waarde zal de snelheid en de doeltreffendheid van het algoritme bepalen. Voor alle optimalisaties is dx vast op 50 nm (25 nm langs elke kant van een sectie). Een eerste optimalisatie met een startlengte van 66 µm en a_s=0.9 slaagt er niet in de efficiëntie duidelijk te verbeteren, doordat de lengte teveel verandert in een verkortingsstap en de goede veranderingen daardoor niet meegenomen worden. Een andere waarde voor a_s, namelijk 0.95, is gebruikt in een tweede optimalisatie, die in figuur 15(a) getoond wordt. De andere gebruikte waarden zijn 31.32 µm als startlengte en een maximum van 20 links-rechts-bewegingen. In tegenstelling tot de eerste poging, worden er hier wel duidelijke verbeteringen vastgesteld die een lengtewijziging ook kunnen overleven. In het lengtegebied waar de transmissie van een parabolische omvormer zakt van ongeveer 80% naar bijna 0%, behoudt de gewijzigde parabool een efficiëntie groter dan 95%, waardoor de lengte van een goed functionerende structuur gehalveerd wordt van 36 (parabool) naar 18 µm (gewijzigde parabool).



(a) Succesvolle optimalisatie (b) Geoptimaliseerde waarden

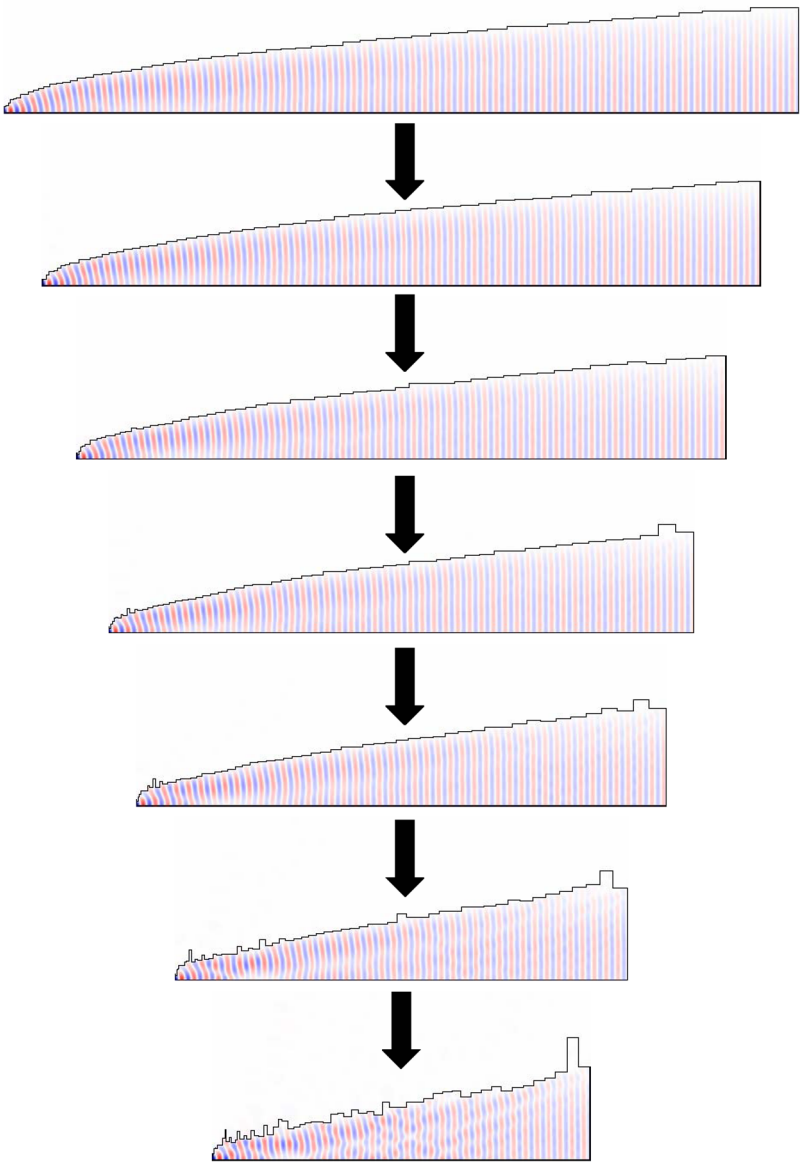
Figuur 15: Links: Optimalisatie met a_s=0.95, maximaal 20 links-rechts-bewegingen en een vertreklengte van 31.32 µm . **Rechts:** Hoogste efficiëntie waarde voor elke gegenereerde omvormerlengte.

Nog andere optimalisaties zijn uitgevoerd met andere parameters. Voor elke omvormerlengte die daarbij bekeken werd, wordt de hoogst verkregen efficiëntie getoond in figuur 15(b). Voor lengtes groter dan 15 µm wordt er een grote efficiëntie bereikt, maar beneden deze lengte zakken de waarden plotseling. We geloven dat deze plotse daling niet intrinsiek is aan het bestudeerde probleem, maar dat het voor kortere lengtes alsmaar moeilijker wordt om een goed lokaal optimum te bereiken. Met waardes voor a_s nog dichter bij een, en met meer links-rechts-bewegingen, zal de efficiëntie ook toenemen voor kortere lengtes. Hierdoor neemt de benodigde rekentijd ook sterk toe, wat onpraktisch is en waardoor we dit niet gecontroleerd hebben.

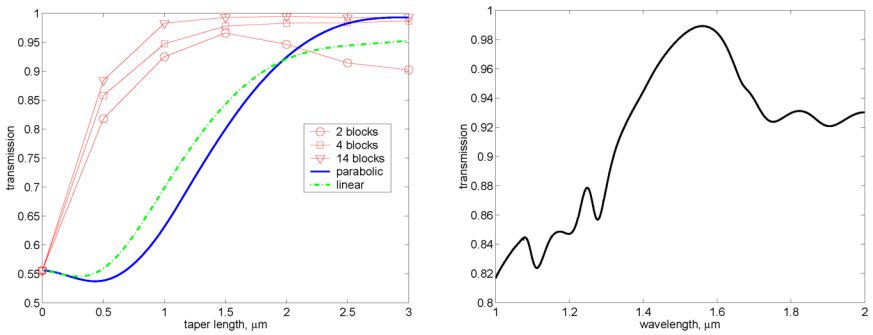
Figuur 16 toont de veldwaarden van sommige geoptimaliseerde configuraties, allen tijdens dezelfde optimalisatie verkregen. Het ontstaan van enkele kenmerkende structuurtjes is duidelijk zichtbaar, zoals de bobbel nabij het brede uiteinde en de verzameling kleine staafjes aan het andere uiteinde. Het vlakke fasefront aan het einde van de structuur is een visuele hint voor een goede efficiëntie.

4.3 Ab Initio Blokomvormers

In plaats van te starten vanuit een gekende en werkende structuur zoals hierboven, gebeurt de optimalisatie van de verschillende golfgeleidersecties hier zonder enige vorm van voorkennis. Om de grootte van de parameterruimte, waarin we nu moeten zoeken naar optima, te illus-



Figuur 16: 7 veldplots (H_y) van geoptimaliseerde omvormers, enkel de bovenste helft is getoond, in- en uitgangsgolfgeleiders niet getoond. Van boven naar beneden zijn de lengtes 37.0, 33.4, 30.2, 27.3, 24.6, 21.1 en 17.7 µm.



(a) Geoptimaliseerde efficiënties (b) Transmissiespectrum

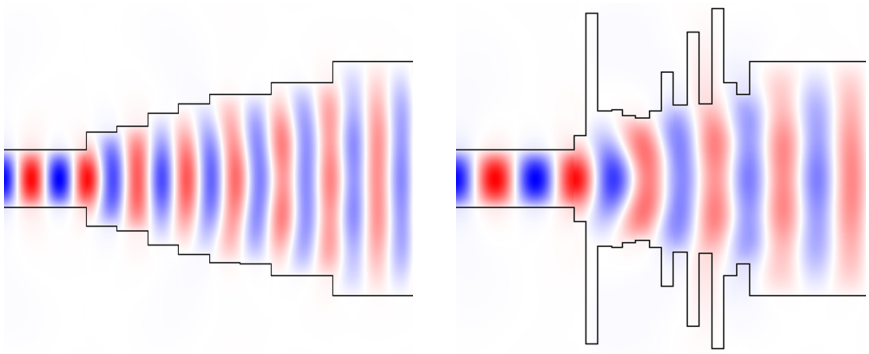
Figuur 17: Links: Efficiëntie als functie van de lengte voor een verscheiden aantal secties. **Rechts:** Transmissiespectrum van de omvormer in figuur 18(b)

treren beschouwen we een omvormer bestaande uit 14 secties met een variabele breedte maar een vaste lengte en laten we de breedtes toe om te variëren tussen 0.3 en 3.0 µm in stapjes van 10 nm. Hierdoor ontstaan er $270^{15} \approx 3 \cdot 10^{36}$ mogelijke combinaties. Om in een dergelijke gigantische multidimensionale parameterruimte naar optima te zoeken, zijn genetische algoritmes uitermate geschikt.

4.3.1 Van 0.5 naar 2.0 µm

Vooreerst wordt de koppeling tussen SOI-golfgeleiders met breedtes van 0.5 en 2.0 µm bestudeerd bij een golflengte van 1.55 µm en a.d.h.v. de effectieve-indexmethode. De optimalisatie wordt herhaald voor verscheidene omvormerlengtes L en voor een verschillend aantal secties N. De lengte van de secties is altijd vast, namelijk $L_i = L/N$. Breedtes mogen variëren tussen 0.3 en 3.0 µm in stapjes van Δx=10 nm.

Als start worden er 250 individuen willekeurig gegenereerd maar enkel de 100 beste vormen samen de eerste generatie. Om 2 individuen te selecteren wordt er roulettewielselectie gebruikt met als fitness de grondmode-transmissie. Uitwisseling van parameters gebeurt door een uniforme crossover en de mutatie van het nageslacht is een gediscretiseerde gaussiaanse mutatie met σ=W_i en μ=$3/4 \cdot \Delta x$. De optimalisatie wordt gestopt van zodra het verschil in fitness tussen het beste en het slechtste individu van de laatste generatie minder dan een half procent s, waarvoor nooit meer dan 1200 generaties nodig waren.



(a) N=8, L=3.0 µm, T=0.99 (b) N=14, L=1.5 µm, T=0.99

Figuur 18: Veldplot (reëel deel van H_y) van twee structuren. Blokomvormers berekend in 2-D voor koppeling tussen 0.5 en 2.0 µm.

Voor elke omvormerlengte wordt het aantal secties gevarieerd van 2 tot 14 in stapjes van 2. Figuur 17(a) toont enkele resultaten, waarbij elk curvepunt het eindresultaat is van een afzonderlijke genetische optimalisatie. Terwijl een gladde lineaire en parabolische omvormer van 1.5 µm lang een respectievelijke efficiëntie hebben van 84% en 80%, laat de structuur met 14 secties bij gelijke lengte meer dan 99% van het inkomende licht door. Zoals te verwachten valt, verhoogt de efficiëntie als het aantal vrijheidsgraden van een structuur toeneemt, wat in dit geval meer secties betekent.

In figuur 18 staan enkele van de resulterende structuren, waarbij het reëel deel van het transversale (y-richting) magnetische veld wordt getoond wanneer er van links met de grondmode wordt geëxciteerd. In figuur 18(a) is de omhullende van de structuur sterk gelijkend op een parabolische omvormer. Door de best gekende omvormer bij deze lengte te reconstrueren wordt de doeltreffendheid van de optimalisatie aangetoond. De omvormer met 14 secties lijkt hoogst onregelmatig maar slaagt erin om de efficiëntie te verhogen.

Men zou kunnen verwachten dat de omvormer in figuur 18(b) een sterk resonant gedrag zou vertoont door de talrijke overgang tussen lage en hoge brekingsindex en de bijhorende reflecties. Wanneer we echter het spectrum van deze structuur bekijken in figuur 17(b) toont dit niet de sterke fluctuaties die typisch zijn voor een resonantie. Voor een relatief breed golflengtegebied (1.5-1.6 µm) is de efficiëntie zelfs hoog (meer dan 98%) en vrij vlak.

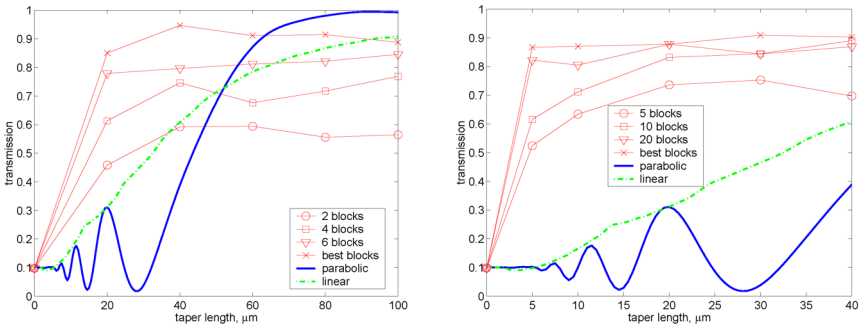
Het is belangrijk om op te merken dat een omvormer tussen twee SOI-golfgeleiders met de bovenstaande breedtes (0.5 en 2.0 µm) reeds adiabatisch is als hij langer is dan 10 µm (zowel lineair als parabolisch). De grootte van een dergelijke omvormer is dus bescheiden en de te winnen chipoppervlakte bijna verwaarloosbaar. De bovenstaande studie moet dus vooral als een voorbeeld gezien worden dat in andere situaties kan herbruikt worden waar meer winst mogelijk is.

4.3.2 Van 0.5 naar 12.0 µm

Een meer praktische situatie is de verbinding tussen golfgeleiders met een breedte van 0.5 µm en 12 µm die zich voordoet als een verticale vezelkoppeling, zoals in figuur 10, met een monomodale SOI-golfgeleider moet verbonden worden.

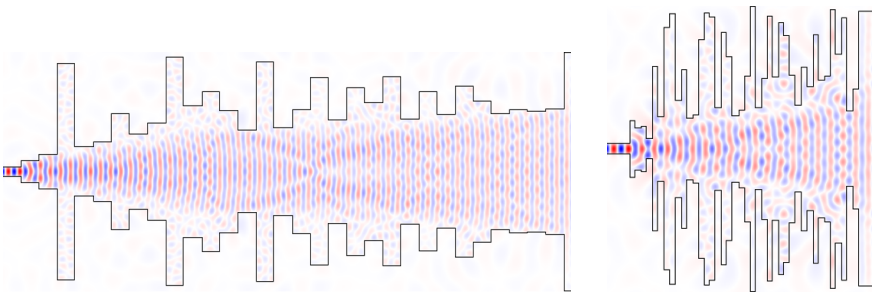
Hetzelfde genetische algoritme als in de smallere koppelsituatie is gebruikt met uitzondering van de toegestane breedtes, die tussen 0.3 en 13.0 µm liggen en kunnen variëren in stapjes van Δx=25 nm. In een eerste reeks optimalisaties wordt de lengte van de omvormers veranderd tussen 0 en 100 µm in stappen van 20 µm, en het aantal secties tussen 2 en 14 in stappen van 2. Een selectie van de resultaten is te zien in figuur 19(a) samen met de transmissie van lineaire en parabolische omvormers. Als de complexiteit van een blokomvormer verhoogt (meer secties) gaat ook de efficiëntie omhoog. Dit is echter niet altijd strikt het geval omdat een genetisch algoritme geen globaal optimum garandeert. De lijn aangeduid met 'best blocks' verbindt de hoogste efficiënties, die afhankelijk van de lengte verschijnen bij 10, 12 of 14 secties. Voor lengtes langer dan 60 µm presteren blokomvormers (tot 14 secties) slechter dan conventionele parabolen.

Een tweede reeks optimalisaties is meer gericht op het lengtegebied waar een grote winst in efficiëntie kan bereikt worden en gebruikt meer secties in een omvormer. Figuur 19(b) toont de resultaten waar 'best blocks' opnieuw de hoogste efficiëntie aanduidt, onafhankelijk van het aantal gebruikte secties. De figuur maakt ook duidelijk dat, behalve voor zeer korte lengtes, de efficiëntie min of meer onafhankelijk is van de lengte maar zeer sterk afhangt van het aantal gebruikte secties. Blokomvormers met slechts enkele secties kunnen een sterke winst geven voor korte lengtes, maar zeer hoge efficiënties, hoger dan 95% (-0.2dB), hebben we in deze koppelsituatie niet behaald.



(a) Grove reeks (b) Fijnere reeks

Figuur 19: Geoptimaliseerde efficiënties voor blokomvormers met een verscheiden aantal secties tussen golfgeleiders met breedtes van 0.5 en 12 µm .



(a) N=30,L=30µm,T=0.91 (b) N=40,L=10µm,T=0.87

Figuur 20: Veldplot (reëel deel van H_y) voor enkele structuren langs links geëxciteerd. Blokomvormers voor koppeling tussen 0.5 en 12.0 µm.

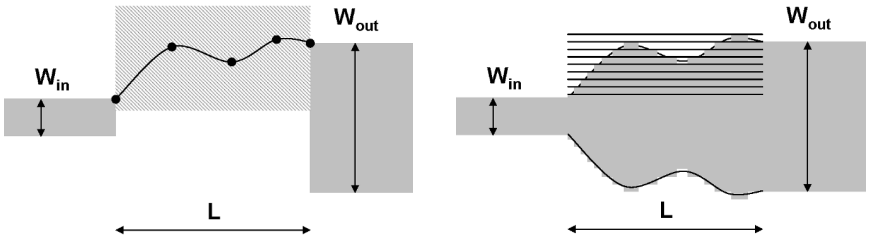
De velden binnenin enkele representatieve structuren zijn te zien in figuur 20. Hoe meer secties een omvormer bevat, hoe grilliger hij ook wordt.

4.4 Spline-Omvormers

Een derde aanpak om de verschillende secties (W_i,L_i) in figuur 13 te optimaliseren wordt in deze paragraaf besproken. In plaats van de parameters rechtstreeks te veranderen zoals hierboven, wordt er nu een omhullende functie geoptimaliseerd, waaruit achteraf de parameters voor elke sectie afgeleid worden.

Beschouw een omvormer met een vaste lengte L tussen twee golfgeleiders met gegeven breedtes W_{in} en W_{uit} zoals in figuur 21 en beschouw slechts een helft afgebakend door de propagatie/spiegelas. Een virtueel toegestaan gebied wordt bepaald tussen de twee golfgeleiders. De onderkant van het gebied zal bepalen hoe breed de smalst toegelaten golfgeleider is en de bovenkant bepaalt de maximale breedte, die het uiteindelijk het totale simulatiegebied en de simulatieduur beïnvloedt. Binnenin dit gebied worden er willekeurig een aantal N punten (z_i, x_i) geplaatst, die de uiteindelijke vorm van de omvormers zullen bepalen. Deze punten worden met elkaar verbonden, in volgorde van stijgende z-waarde, met behulp van een derdegraadsspline. Dit is een stuksgewijze derdegraadsveelterm die voldoet aan de voorwaarde dat in elk oorspronkelijk punt de nulde en de eerste afgeleide continu zijn. Voor de binnenste punten bepaalt deze voorwaarde de spline reeds volledig. Maar voor de buitenste punten moet er een extra voorwaarde gesteld worden, met name dat de tweede afgeleide nul wordt, wat het voordeel biedt dat een spline gebaseerd op nul punten samenvalt met een lineaire omvormer. Een derdegraadsspline met deze extra randvoorwaarde wordt een natuurlijke derdegraadsspline genoemd. Een mogelijke andere randvoorwaarde is het stellen van een zekere waarde voor de eerste afgeleide in de buitenste punten. In dat geval wordt de spline een geklemde derdegraadsspline genoemd. De keuze voor een welbepaalde randvoorwaarde sluit echter bepaalde structuren niet bij voorbaat uit aangezien een extra punt dichtbij het buitenste punt de voorwaarde in dit buitenste punt onbelangrijk maakt. Als we over het aantal spline punten spreken, laten we hierna steeds de 2 buitenste punten buiten beschouwing.

Met deze continue kromme worden er dan z-invariante golfgeleiders gedefinieerd, nodig om de structuur met een EEM te berekenen.



(a) Willekeurige punten in toegestaan gebied

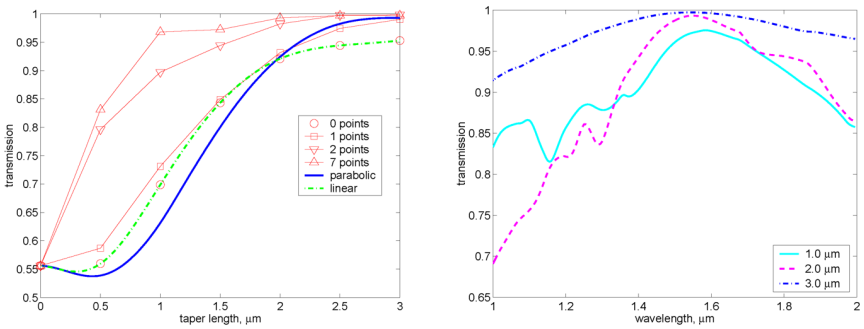
(b) Verbonden curve en discretisatie

Figuur 21: Constructie van een omvormer vertrekkend van willekeurig geplaatste punten, de kromme die ze verbindt en de discretisatie via een vast x-rooster.

Verschillende discretisatieschema's kunnen bedacht worden, zoals een vaste sectielengte of een vaste stap in de breedterichting. Alhoewel de eerste het gemakkelijkst wiskundig valt te implementeren, is er voor de tweede gekozen omdat die veel voordelen biedt in combinatie met eigenmode-expansie. Gebaseerd op de vastgelegde breedtestap Δx en W_{in} wordt er een rooster in de x-richting vastgelegd dat de toegelaten breedtes zal bepalen, zie figuur 21(b). Door te kijken waar de spline halfweg tussen twee roosterpunten passeert, worden de lengtes van de secties bepaald. Samen met enkele extra regeltjes rond lokale optima, legt deze procedure eenduidig vast hoe een spline omgezet wordt naar een gediscretiseerde structuur, klaar om berekend te worden.

4.4.1 Van 0.5 naar 2.0 µm

Het genetisch algoritme volgt hetzelfde stramien als normaal behalve voor de crossover, waar nu een eenpuntscrossover gebruikt is. Gedurende deze crossover blijven de twee coördinaten van een punt samen. Elk punt wordt gemuteerd met dezelfde gaussiaanse mutatie-operator als hierboven waarbij σ=7.5 nm en er niet gediscretiseerd wordt. Nadien worden alle punten van een nieuw individu gerangschikt volgens stijgende z-coördinaat om te garanderen dat de spline een functie is. Zelfs als alle punten binnen het toegestane gebied liggen kan de spline er nog uitsteken. Als dit het geval is, wordt het nageslacht geweigerd en de crossover en mutatie opnieuw uitgevoerd totdat er een geldige splinecurve is. De discretisatie van de splinekromme is gebaseerd op Δx= 12.5 nm, wat wil zeggen dat twee opeenvolgende golfgeleidersecties 25 nm verschillen qua breedte.

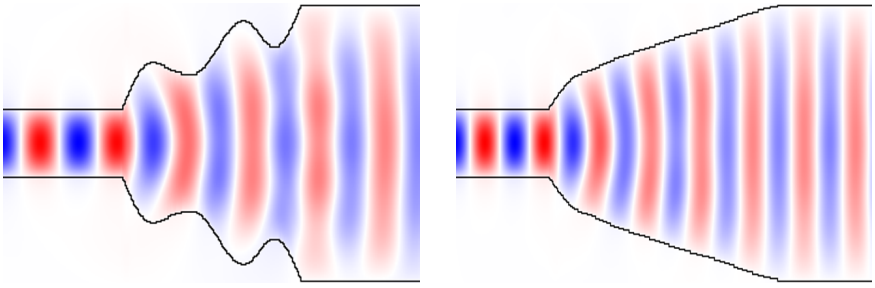


(a) Geoptimaliseerde efficiënties (b) Enkele spectra

Figuur 22: Links: Geoptimaliseerde efficiënties van spline omvormers, het aantal punten is aangeduid in de legende. **Rechts:** Transmissiespectra voor spline omvormers met 7 punten, de lengte is aangeduid in de legende.

In- en uitgangsgolfgeleiders hebben breedtes van 0.5 en 2.0 en de lengte van de omvormers gaat van 0.5 tot 3 µm. Voor elke lengte wordt het aantal punten gevarieerd van 0 tot 7. Elk punt in de figuur 22(a) is het eindresultaat van een afzonderlijke optimalisatie, behalve voor de kromme gebaseerd op 0 punten waar er niets te optimaliseren valt. Voor drie van de structuren met 7 punten is het spectrum getoond in figuur 22(b), waar het duidelijk is dat de bandbreedte kleiner wordt als lengte afneemt. In de figuur hebben niet alle spectra hun maximum bij de golflengte van 1.55 µm, waarbij er geoptimaliseerd is. Dit toont nog eens aan dat het genetisch algoritme niet het globale optimum kan identificeren. Toch hebben we 2 redenen om te geloven dat onze structuren dit globale optimum vrij dicht naderen. Een eerste reden is dat de langere omvormers benaderend parabolisch zijn, zie figuur 23(b) als voorbeeld. Dit kan je verwachten als de lengtes de voorwaarde voor een adiabatisch gedrag benaderen.

De tweede is de convergentie binnen de resultaten. Voor een vaste lengte maar een stijgend aantal punten, veranderen structuren niet drastisch meer van vorm. De veldplots die dit illustreren zijn te vinden als illustratie bij de Engelse tekst. Aangezien verschillende structuren het gevolg zijn van onafhankelijke optimalisaties besluiten we dat deze structuren convergeren naar een optimale vorm die door een groot aantal punten beschreven wordt.



(a) N=7,L=1.5µm,T=0.972 (b) N=7,L=2.5µm,T=0.997

Figuur 23: Veldwaarden van enkele spline omvormers voor de koppeling tussen 0.5 en 2.0 µm. Figuren zijn niet op dezelfde schaal.

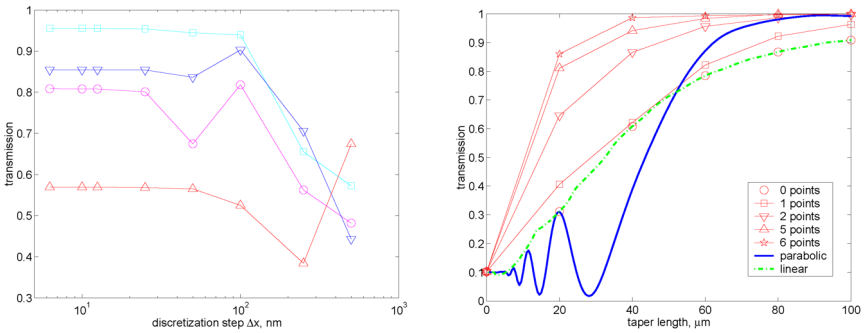
4.4.2 Van 0.5 naar 12.0 µm

Hetzelfde genetisch algoritme als hierboven is gebruikt om deze koppeling te optimaliseren, behalve voor de discretisatiestap Δx die hier is verdubbeld tot 25 nm (50 nm verschil tussen twee opeenvolgende breedtes). Om de invloed van deze stap op de efficiëntie te bestuderen, is hij gevarieerd tussen 6 en 500 nm voor 4 willekeurig gekozen spline omvormers. In figuur 24(a) is het duidelijk dat er een convergentie is voor kleine stappen, met 25 nm als bovenlimiet en daarboven een onstabiel gebied.

Omvormerlengtes variëren van 0 tot 100 µm in stappen van 20 µm en de splines zijn gebaseerd op 0 tot 7 punten. De resultaten zijn te zien in figuur 24(b) en een bijhorende veldplot in figuur 25.

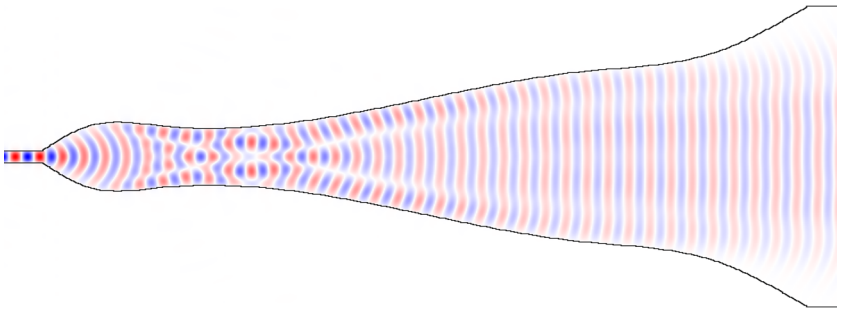
4.5 Andere Optimalisatiefuncties

Tot nu toe hebben we altijd dezelfde eigenschap van een omvormer geoptimaliseerd, namelijk de vermogentransmissie van de grondmode. Door een andere functie te gaan optimaliseren, zal de resulterende omvormer goed zijn in het omzetten van inkomende mode i naar uitgaande mode j. Een overzicht van ons werk in dit gebied is beschikbaar in de Engelse tekst. Naast mode-omvormers zoals een omvormer die de grondmode omzet in de eerste of de tweede mode, worden er ook omvormers ontworpen die efficiënt in en uit een fotonischekristalgolfgeleider koppelen.



(a) Efficiëntie versus Δx (b) Geoptimaliseerde efficiënties

Figuur 24: Links: Efficiëntieconvergentie voor kleine discretisatiestappen. **Rechts:** Geoptimaliseerde efficiënties voor spline omvormers van verschillende lengte. Het aantal punten is aangeduid in de legende. Koppeling van 0.5 naar 12.0 µm.



Figuur 25: N=4,L=20µm,T=0.90. Veldwaarden voor een spline omvormer die koppelt tussen 0.5 en 12.0 µm.

5. Metingen

Om sommige van de hiervoor verkregen ontwerpen te verifiëren, werden ze verwezenlijkt in een optische chip. Als materiaalsysteem werd silicium-op-isolator verkozen, vanwege de vertrouwdheid van onze onderzoeksgroep ermee en omdat we daarvoor de CMOS-faciliteiten in IMEC, Leuven kunnen aanwenden.

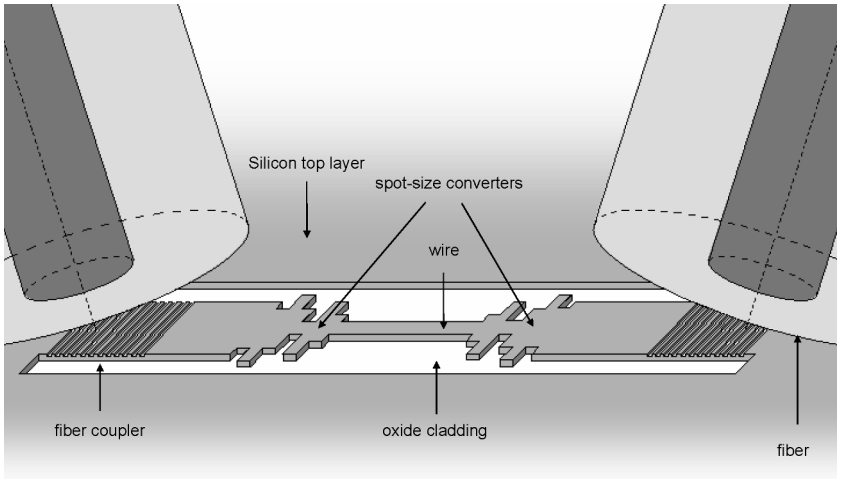
5.1 Fabricage

Voor het verloop en de details van de fabricage verwijzen we graag de Engelse tekst. Als patroondefinitie is er gekozen voor diep-UV-lithografie bij 248 nm. De gebruikelijke processtappen in CMOS-fabricage zijn zoveel mogelijk herbruikt. Na de processing in Leuven komen de volledige wafers naar Gent, waar ze in bruikbaardere groottes gekliefd worden. Door het gebruik van verticale-vezelkoppeling als meetwijze, is dit klieven niet zeer kritiek en moet er ook niet verdund of gepolijst worden. De uiteindelijke uit te meten stalen zijn ongeveer 3.0x1.5 cm^2 en bevatten 6 keer hetzelfde patroon, maar soms met een andere dosis gefabriceerd, waardoor de afmetingen een klein beetje verschillen. Hierdoor worden eventuele consequente afwijkingen in de fabricage opgevangen.

5.2 Meetopstelling

Voor alle metingen is dezelfde opstelling gebruikt, zie figuur 26. De te meten omvormers zitten ingewerkt in een grotere structuur die links begint met een verticale-vezelkoppelaar, in essentie een eendimensionaal tweede-orderooster. De golfgeleider, waarin dit rooster is verwerkt, is 10 µm breed en loopt door tot aan de eerste omvormer. Deze heeft als uitgang een draadgolfgeleider van 0.5 µm die doorloopt tot de volgende omvormer. De tweede helft van de structuur vormt het spiegelbeeld van de eerste helft. De roosterkoppelingen zijn ontworpen om in of uit te koppelen naar een vezeltip die 10° schuin staat t.o.v. de normaal op het rooster, waardoor de symmetrie gebroken wordt en reflecties aan het rooster in de brede golfgeleider onderdrukt worden. Een instelbare laser (1500-1640 nm) stuurt licht in de ingangsvezel, met tussendoor nog een polarisatiedraaier. De uitgangsvezel vangt het licht op en stuurt het door naar de detector.

Beide vezeluiteinden en het meetmonster kunnen nauwkeurig gepositioneerd worden t.o.v. elkaar d.m.v. precisieschroeven en een piëzo-



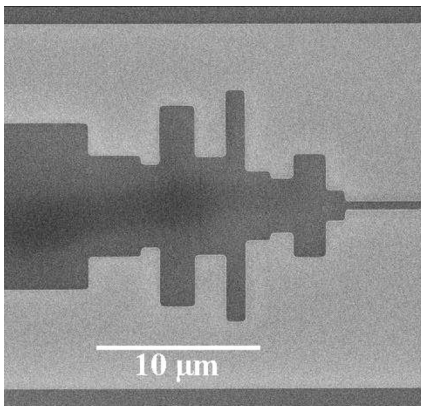
Figuur 26: Meetopstelling m.b.v. verticale-vezelkoppeling. Ingezoemd op de chip en vezeluiteinden.

elektrische sturing. Op alle meetmonsters staan de twee vezelkoppelingen 7 mm uiteen. Structuren verschillen onderling door wat er tussen de twee stukken rechte brede golfgeleider staat. In sommige referentiestructuren loopt deze golfgeleider van 10 µm breed volledig door tussen de twee roosters, in andere verbinden adiabatische lineaire omvormers naar een stukje draadgolfgeleider in het midden.

Hieronder worden er verschillende metingen beschreven. De titel van een paragraaf komt overeen met de naam van het masker waarmee de behandelde structuren definieerd zijn. Enkel de meest succesvolle metingen worden getoond.

5.3 PICCO-03

Om het verlies van een enkele omvormer te kunnen kwantificeren, moeten alle andere mogelijke verliezen vermeden worden of goed gekend zijn. Daarvoor zijn er referentiestructuren voorzien die gelijk zijn aan de structuren in figuur 26, behalve dat de compacte omvormer is vervangen door een adiabatische lineaire omvormer. Toch zijn er nog wat problemen opgedoken rond de normalisatie van de metingen en het kwantificeren van de verliezen. Een uitvoerige beschrijving is voorhanden in de Engelse tekst. De belangrijke conclusie is dat de kleinste fout op de verliezen gemeten wordt bij de golflengte waar de vezelkoppeling het efficiënst is, en dat voor andere golflengtes deze fout zeer snel

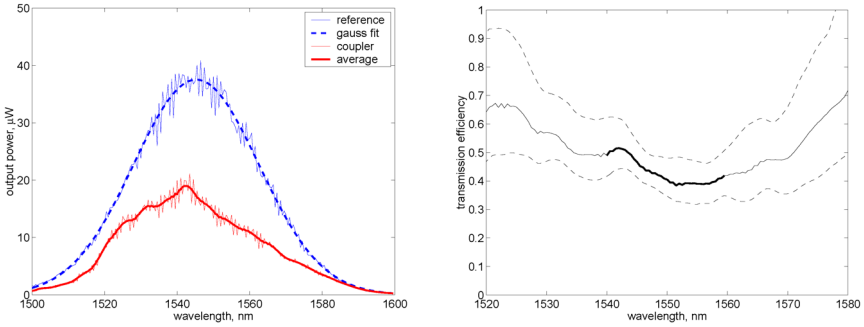


Figuur 27: SEM-beeld van de component die koppelt van 10.0 naar 0.56 µm op een afstand van 15.4 µm. Lichtgrijs komt overeen met de blootgeëtste laag silica (glas).

oploopt. Daarom wordt, in wat volgt, bijna alle dataverwerking verricht bij deze optimale golflengte, die verschilt van sample tot sample. Figuur 27 toont een SEM (scanning electron microscope) beeld van de beste component van deze reeks. Oorspronkelijk is er in 2-D een blokomvormer geoptimaliseerd met 10 golfgeleidersecties, die in breedtes varieerden tussen 0.3 en 13.0 µm en in lengte tussen 0.3 en 3.0 µm. De resulterende structuur had een vermogenstransmissie van 78% in 2-D, maar slechts 71% in een rigoreuze 3-D berekening. Een verder optimalisatie in 3-D, met de lokale optimalisatiemode van Kallistos, heeft deze transmissie ietwat verhoogd tot 72%.

Op het meetmonster zit deze structuur ingewerkt in een configuratie zoals in figuur 26. Het transmissiespectrum van de beste omvormerconfiguratie is te zien in figuur 28(a) samen met de transmissie van de dichtsbijzijnde referentiestructuur.

Het opgemeten spectrum van de referentiestructuur valt bijna samen met de gaussiaan die eraan gefit is, wat een welbekend feit is uit de literatuur rond verticale-vezelkoppeling. De metingen van de structuur met twee blokomvormers zijn onregelmatiger en daar wordt een spectraal gemiddelde over 9 punten getoond. Om te normaliseren moeten deze twee curves door elkaar gedeeld worden, wat grote fouten in de uiteindelijke efficiëntiewaarden introduceert, zoals reeds hierboven vermeld. De verkregen efficiëntiecurve wordt samen met een onderen een bovengrens voor de fout getoond in 28(b).



(a) Opgemeten spectra (b) Genormaliseerde efficiëntie

Figuur 28: Links: gemeten spectra van een referentiestructuur en een blokomvormer met een gaussiaanse fit een spectraal gemiddelde. **Rechts:** Efficiëntiespectrum (volle lijn) en onder- en bovengrens voor de fout (stippellijn).

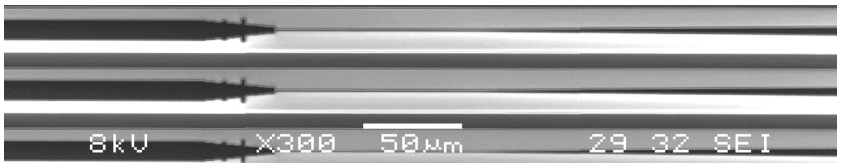
In het gebied waar de fout minimaal is, ligt de efficiëntie tussen 40% en 52%. Aangezien deze waarde 2 blokomvormers bevat, heeft een enkele omvormer een efficiëntie tussen 63% en 72%, wat vrij goed overeenkomt met de berekende waarde voor deze component.

5.4 PICCO-04

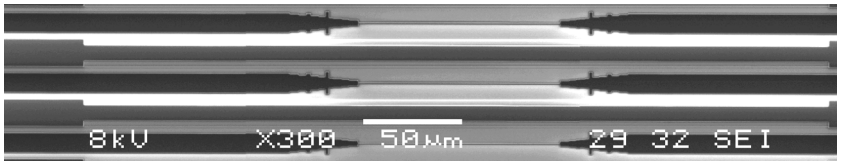
De structuren die met het PICCO-04 masker gerealiseerd zijn, laten een meer systematische studie toe van de performantie van blokomvormers in vergelijking met lineaire en parabolische. Figuur 29 toont SEM-beelden van enkele structuren. Dezelfde blokomvormer is telkens 7 keer herhaald met een licht verschillende vorm in een poging om onder- of overbelichting tussen de fabricage op te vangen. Aangezien de onveranderde vorm ook bijna altijd de beste is gebleken, was dez compensatie achteraf gezien niet nodig.

Voor deze samples ligt het maximum van de vezelkoppeling bij ongeveer 1535 nm. Deze golflengte is na een eerste identificatie ook altijd gebruikt als alignatiegolflengte.

Figuur 30 combineert alle PICCO-04 metingen op lineaire en parabolische omvormers en omvormers met 5 of 10 blokken. De lineaire en parabolische omvormers tonen een lengtegedrag dat zeer gelijkend is aan de theoretische waarden die in een vorige paragraaf bekomen werden. De efficiëntie van de blokomvormers ligt hoger dan die voor lineaire en parabolische van gelijke lengte. De 5-blokomvormer met een

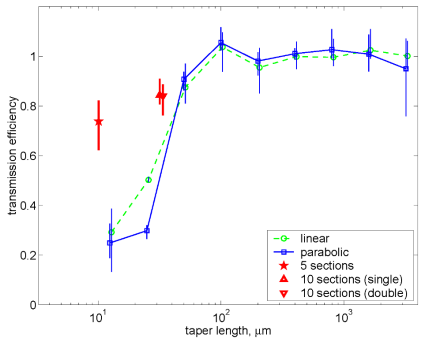


(a) Component met een blok- en een lineaire omvormer

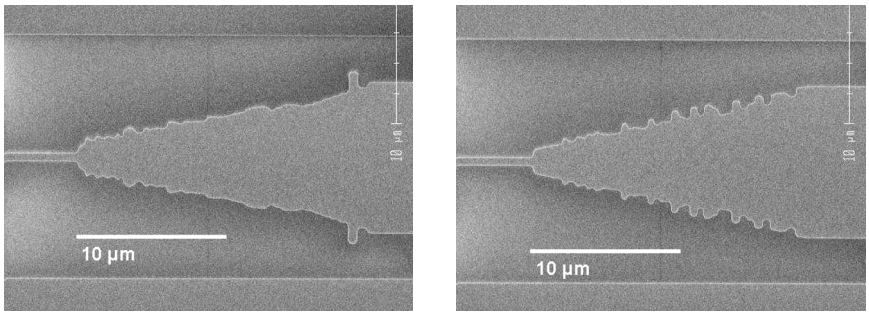


(b) Component met twee symmetrische blokomvormers

Figuur 29: Componenten gefabriceerd a.d.h.v. het PICCO-04 masker. De lengteschaal is aangeduid in de figuur.



Figuur 30: Gemeten efficiëntiewaarden van sommige PICCO-04 structuren: lineaire en parabolische omvormers met lengtes van 12.5,25,50,. . .,3200 µm en blokomvormers met 5 (10.1 µm) en 10 (33.8 µm) secties.



(a) Gewijzigde parabool (19.1 µm) (b) Gewijzigde parabool (17.8 µm)

Figuur 31: SEM-beelden van gewijzigde parabolische omvormers op het PICCO-01 masker.

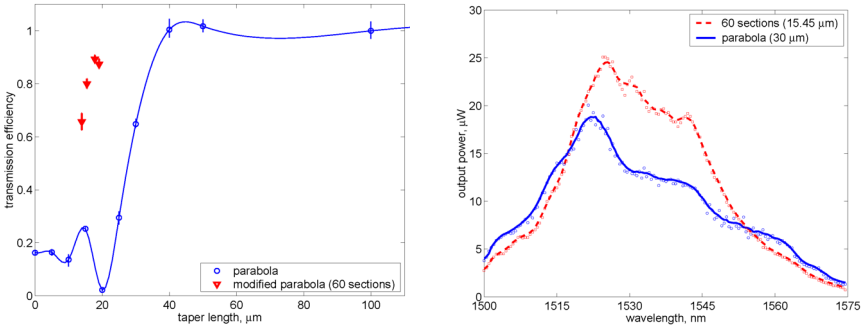
lengte van 10.1 µm laat 74% van het inkomende licht door en de 10-blokomvormer met een lengte van 33.8% ongeveer 84%. De waardes berekend in 3-D zijn respectievelijk 70% en 79% en liggen binnen het foutinterval. Het is ook geruststelllend om te zien dat de metingen op de enkele en de dubbele structuren, zie figuur 29, tot gelijkaardige waarden leiden.

5.5 PICMOS-01

In de laatste reeks structuren zijn er niet alleen blokomvormers maar ook gewijzigde parabolen gerealiseerd. Figuur 31 toont SEM-beelden van enkele van deze structuren. De breedte van de draadgolfgeleiders is hier 0.5 in plaats van 0.56 µm. Parabolische omvormers van verscheidene lengtes zijn ook aanwezig met deze keer meer beschikbare lengtes in het interessante gebied tussen 0 en 200 µm.

Figuur 32(a) bundelt de resultaten van de gewijzigde en ongewijzigde parabolen. Voor deze reeks metingen ligt de piek van de vezelkoppeling bij 1532 nm. Vier verschillende gewijzigde parabolen zijn aanwezig, met als lengtes 13.9, 15.45, 17.8 en 19.07 µm en als bijhorende 3-D efficiënties 62, 69, 82 en 80%. De gemeten waarden liggen iets hoger dan de berekende, maar dezelfde trend is duidelijk zichtbaar.

In figuur 32(b) worden de opgemeten uitgangsvermogenspectra getoond voor twee van de structuren uit figuur 32(a). Het ingangsvermogen is in beide gevallen 1 mW en omdat het vermogen wordt gemeten van vezeltip naar vezeltip en daarbij twee omvormers passeert, moet er na normalisatie nog een vierkantswortel getrokken worden om de



(a) Efficiënties van gewijzigde parabolische omvormers

(b) Twee spectra (uitgangsvermogen) en het spectraal gemiddelde van 9 punten

Figuur 32: Metingen op PICMOS-01 structuren, zowel gewijzigde als ongewijzigde parabolische omvormers voor koppeling van 0.5 naar 10 µm.

efficiëntie van een enkele omvormer te bekomen. Waar de vezelkoppeling groot is, presteert de gewijzigde parabool duidelijk beter dan een ongewijzigde parabool die bijna tweemaal zo lang is.

6. Algemeen Besluit

Hoge verpakkingskosten en de stijgende complexiteit van systemen voor optische communicatie, zijn de drijvende kracht achter optische integratie. Het einddoel van die integratie is de realisatie van fotonische geïntegreerde circuits. Binnenin dergelijke circuits kunnen verschillende types en groottes van golfgeleiders samen bestaan en lichtbundelomvormers zullen nodig zijn om die onderling met elkaar te verbinden. Om dure plaats op de chip te besparen moeten die omvormers zo klein mogelijk gemaakt worden.

Uit de literatuur rond omvormers is het welbekend dat een parabool de vorm is die de koppeling tussen de grondmode en hogereordemodes minimaliseert en daardoor tot de kortste adiabatische omvormers leidt. Om driedimensionale omvormers te realiseren, moet er in beide dimensies een parabolische vorm voorzien worden wat technologisch geen evidentie is. Daarom zijn er reeds verscheidene nieuwe vormen voorgesteld die het verbreden van de bundel in de verticale richting behandelen. De toepassing van deze nieuwe ideeën in de hori-

zontale richting opent de deur voor 2-D omvormers die kleiner zijn dan conventionele.

Bij het ontwerp van micro-optische componenten is het dikwijls nuttig om niet alles exact in 3-D te berekenen maar om een optisch probleem te gaan benaderen door een analoog 2-D probleem d.m.v. een effectieve-indextransformatie. Dit laat snellere berekeningen toe wat dan weer structurele optimalisaties van componenten mogelijk maakt. Genetische algoritmes zijn goed geschikt om deze problemen aan te pakken en zijn wijd verspreid in elektromagnetische optimalisaties.

Drie verschillende manieren om een omvormer te construeren zijn bestudeerd in dit werk, een gebaseerd op het wijzigen van een parabolische omvormer, en twee op ab-initio optimalisaties. Daarvan werkt er een rechtstreeks in op de parameters van een reeks golfgeleidersecties en het andere optimaliseert de vorm van de omhullende van deze secties. De drie methodes slagen erin om de minimum omvormerlengte waarbij er een hoge transmissie wordt bereikt te halveren.

Een selectie van deze componenten is gerealiseerd in SOI en uitgemeten aan de hand van verticale vezelkoppeling, waarbij er wat extra aandacht vereist is tijdens de dataverwerking. De drie uitgemeten reeksen tonen aan dat theorie en experiment vrij goed overeenstemmen. In het bijzonder, zijn er binnen elke reeks geoptimaliseerde structuren gevonden die beter presteren dan lineaire en parabolische omvormers van gelijke lengte. Omdat de gemeten efficiënties echter nog onder de waarden liggen, die aanvaardbaar zijn in commerciële componenten, is verder onderzoek zeker nog vereist.

English Text

Chapter 1

Introduction

1.1 Telecommunication

1.1.1 Some History

With the invention of the electric telegraph around 1830, humanity disposed for the first time of a reliable form of communication that was faster than a person could travel. Earlier mechanical attempts like the *Chappe Semaphore* [1] in France could also achieve high speeds (Paris - Lille in under an hour) but never functioned well during the night and in heavy weather. Moreover crossing an ocean would always be out of reach of these systems but became reality for the telegraph in 1866 with the construction of the transoceanic *Anglo-American Telegraph Company* cable.

Over the years electric communication was improved and in 1884 Ernst Siemens patented the coaxial cable in Germany. It would take more than 70 years however, until 1956, for this discovery to end up on the ocean floor in the form of the first transatlantic telephone cable *TAT-1*, which could initially carry 36 conversations simultaneously. Until then intercontinental telephony had only been possible using radio communication.

In the meantime, the search for communication media with lower absorption (for longer distances) and higher bandwidth (for higher throughput) had shifted from electrical to optical systems. In 1970 researchers at *Corning Glass Works* had demonstrated the first useful optical fiber with 17 dB optic attenuation loss per kilometer [2]. This new technology proved its maturity in 1988 with the realization of *TAT-8*,

Figure 1.1: Scheme of an optical link utilizing wavelength-division multiplexing (WDM).

the first transatlantic telephone cable using optical fiber and initially carrying 40.000 telephone calls.

Since then the attenuation of an optical fiber has been lowered to below $0.15\ dB/km$ [3], meaning that after 20 km more than half of the launched light is still in the fiber. Also since then the capacity of long-distance communication lines is no longer counted in a number of telephone calls but in gigabits per second (Gbps), as the internet and the world wide web have been responsible for a spectacular increase in data communication, of which voice traffic now only makes up a fraction.

1.1.2 Optical Integration

To further increase the data stream through an optical fiber, not the fiber medium itself but mainly the machinery at both fiber ends is being improved. A faster on-off switching of the laser light can increase the bandwidth and is accomplished by improving laser diodes, external modulators and detectors. The biggest improvement in recent years, however, has come from wavelength division multiplexing or WDM, a technique that increases the fiber bandwidth by sending several data streams, each using a different wavelength, through the same fiber, see figure 1.1. As an example the transatlantic *TAT-14* cable, which came into operation in 2000, achieves 160 Gbps by having 16 wavelength channels each modulated at 10 Gbps.

The increased system complexity naturally comes at a cost, in this case as a considerable number of extra components. That way a modern optical link can contain the following components: multiplexers, demultiplexers, laser diodes, isolators, external modulators, semiconductor optical amplifiers (SOA), variable optical attenuators (VOA), optical pumps for Raman amplification, Erbium doped amplifiers (EDFA),

components for dispersion compensation, optical interleavers, regenerators, ...

Standard practice today is to build each component into an individually pigtailed (and often temperature-controlled) package and to connect these components into one big component train. The advantage is a modular system, where a failing component can be easily replaced and where the system builder can choose between components from various vendors. The big disadvantage is the high price of a single component, as the cost of assembly and packaging can contribute for more than 75% to the total component cost.

To overcome this price problem, optical research departments have been working for years to combine the functionality of various optical components into a single one. A few examples of this quest for integration:

- combining a laser diode and an isolator [4]
- multi-wavelength lasers [5][6]
- optical cross-connects [7]
- combining an optical amplifier and a photodiode [8]
- channel selector and receiver [9]

Figure 1.2 shows a state-of-the-art example of a photonic integrated circuit with 9 arrayed waveguide gratings and 40 photodetectors.

Most industry observers believe that optical integration is the way to go and a lot of research is focused on achieving this goal. In which form, however, this integration ultimately will happen still remains to be seen.

Part of the present research specializes in monolithic integration, meaning that all components are built into the same material system, e.g. indium phosphide (InP) or gallium arsenide (GaAs). As active components (light generation, light detection, fast modulation) can so far only be realized in a semiconductor with a direct band gap, monolithic integration schemes combining active and passive (light guiding, slower modulation) components normally make use of InP or GaAs. Because the waveguide property requirements for active and passive components are different, techniques like local regrowth, quantum-well intermixing or polarization rotation are needed to make monolithic integration work. To circumvent the need for a direct band gap several

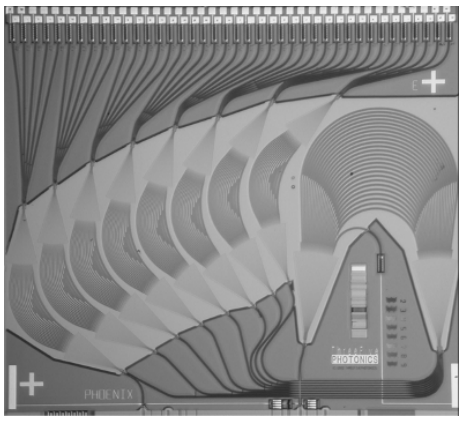


Figure 1.2: Example of state-of-the-art photonic integrated circuit containing 9 AWGs and 40 photodetectors. Courtesy III-V Photonics BV, an ASIP Inc Company.

groups are striving for light emission in an indirect band gap material - mostly the cheap silicon - by nanostructuring, doping with rare-earths [10] or by using Raman amplification [11].

Because monolithic integration demands compromises in component design, other integration techniques are also under investigation. Passive components can be cheaply manufactured in established materials like silica (silica-on-silicon) or silicon (silicon-on-insulator) and then combined with active components manufactured in InP or GaAs, having the best of both worlds. This combining can take different forms, which we will denote with the terms *hybrid* and *heterogeneous*.

In hybrid integration the two components are manufactured separately to an almost final state and then combined. A nice overview can be found in [6]. Examples are the insertion of III-V laser diodes in an etched trench in silicon, or the combination of a silica AWG and a III-V SOA array on a silicon bench to form a multi-wavelength laser, by placing the two in close contact to each other. As the assembly of the various components mostly happens manually, hybrid integration is more a research method than a viable mass market method.

Heterogeneous integration on the other hand could become a mass production technique but is still in its infancy. Here, combining the various materials is one of the first steps of manufacturing and preferably applies wafer scale techniques. The components are then defined in the heterogeneous material system. Examples are the wafer bonding

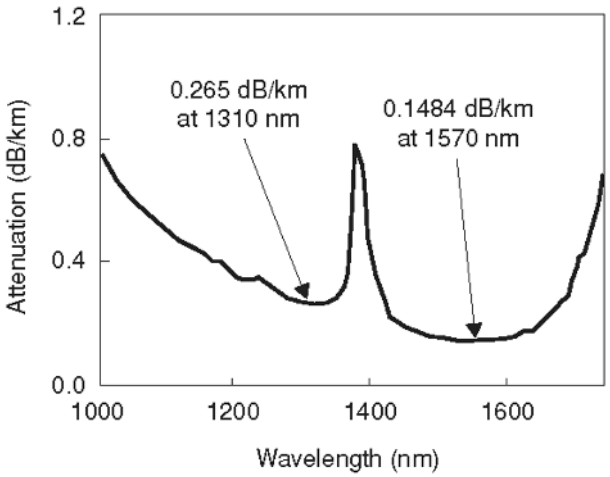


Figure 1.3: Attenuation of the optical fiber described in [3] as a function of wavelength.

of InP stamp sized films on top of silicon using BCB [12] or using the SmartCut technology [13].

Which of the optical integration methods will ultimately lead to commercial components is a topic of active, on-going research. But a cheap commercial platform for optical integration would have a very broad application area.

1.2 Optical Sensing

In section 1.1.1 it was mentioned that the record low attenuation loss of an optical fiber can be as low as $0.15\,dB/km$ [3]. This low loss is limited to a narrow spectral region, as can be seen in figure 1.3. For shorter wavelengths, attenuation losses are higher due to Rayleigh or dipole scattering.

Although an atom (or a molecule) is neutral, as the negative charge of the electron cloud compensates the positive charge of the nucleus (or nuclei), it still has a dipole moment due to the separation of these charges. The resonance frequencies of this dipole lie at shorter wavelengths (typically visible and UV), but even electromagnetic waves with energies too low to excite the dipole at resonance will still weakly excite the dipole, which will then irradiate this energy. Due to inhomogeneities in the solid material not all radiating dipoles will contribute

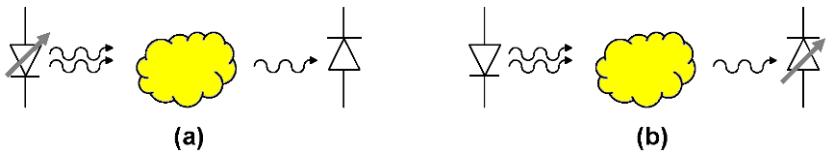


Figure 1.4: Two set-ups for conducting spectroscopic measurements. **(a)** Using a tunable light source. **(a)** Using a tunable light detector.

in the propagation direction of the light and some light will be scattered to other directions. Below resonance the wavelength dependency of Rayleigh scattering goes with λ^{-4}.

Around the wavelength of $1400\ nm$ there is an attenuation peak due to molecular vibrations. Consider a molecular bond that is more ionic than covalent (like in NaCl), in which case it can be regarded as the combination of a positively and a negatively charged atom. When the distance between these atoms changes, so will the molecular dipole moment and again there can be interaction between an electromagnetic wave and the changing dipole, leading to attenuation of the electromagnetic wave. The peak in optical fiber attenuation around $1400\ nm$ is due to vibrations of the O-H or hydroxyl bond, and is often called the *water peak*. As hydrogen is not intrinsic to silica, this peak can be eliminated by ultra-pure processing (no residual hydrogen) or can be shifted by replacing H by D (deuterium) which is heavier and moves the peak to lower energies. Other molecular bonds like Si-O cause attenuation peaks at longer wavelengths and, as the molecules are intrinsic to silica, can not be eliminated.

Molecular vibrations clearly limit which part of the spectrum is useful in combination with optical fiber. But some of the light bi- and triatomic molecules play a very important role in everyday life, like H_2O (weather), CO (unbalanced combustion), CO_2 (greenhouse effect), CH_4 (natural gas, greenhouse effect), SO_2 (acid rain) and NO_x (car exhaust, ozone formation). Due to the molecular vibrations each of these gases has a specific absorption spectrum in the infrared (like a finger print), which can be used to detect it. The research based around this mechanism is called infrared spectroscopy and is indispensable for areas like chemistry, biology and medicine.

A lot of the components developed for optical communication can be adapted for use in spectroscopy, for examples see [14] and [15]. In figure 1.4 two different schematic set-ups to conduct spectroscopy are shown, using a tunable light source or a tunable light detector. The

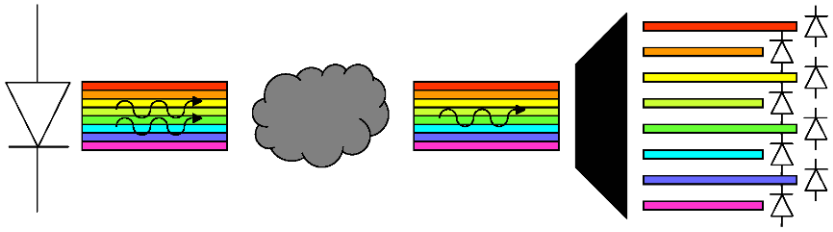


Figure 1.5: Measurement set-up where a WDM detector is applied for gas spectroscopy.

first set-up is nowadays realized using a tungsten filament, serving as a broadband light source and a rotatable diffraction grating, which together form a tunable light source. A second realization is by temperature tuning of a DFB laser, across a small band. A possible scheme for the second set-up is shown in figure 1.5. A broadband light source illuminates the sample under study and the output is sent through a wavelength splitter where each arm has its own detector. A component similar to the one in figure 1.2 could function as the waveguide splitter with integrated detector array.

After the *2000 dot-com bubble* a lot of start-ups that had been developing optical components for telecommunication had to reorient themselves to avoid bankruptcy and are now adapting their research to serve in optical sensing, both for medical as environmental applications.

The exploration of these new application areas together with the rapidly growing market segment of fiber to the home (FTTH) in the established telecommunication market, brings a lot of research and industrial interest into optical integration and photonic chips.

1.3 Guiding Light

Light can be guided by different principles. The best known and conventional way to guide light is by using its property to reside in the material with the highest refractive index, as already demonstrated by John Tyndall, in the nineteenth century, see figure 1.6.

Instead of water as guiding layer and air as surrounding or cladding layer, more advanced combinations are in use today, like doped glass/undoped glass, silicon/glass, silicon/air or quaternary materials like $In_xGa_{1-x}As_yP_{1-y}$ with different (x,y) combinations. This guiding principle is easiest to explain using Snell's law and total internal re-

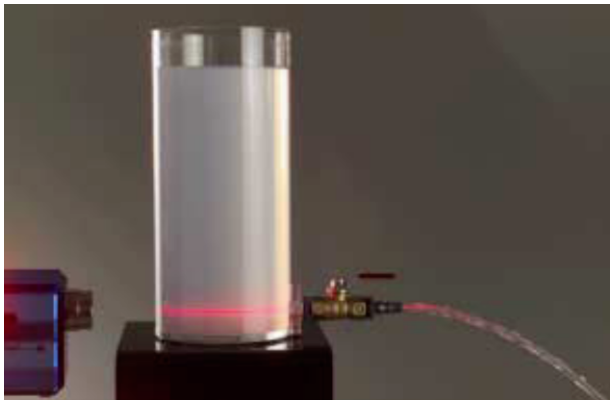


Figure 1.6: Modern version of the original experiment set up by John Tyndall to demonstrate that light can be guided by a water jet.

flection and is therefore known as *guiding by total internal reflection* or shortly *TIR guiding*.

Another way of guiding light is by using mirrors. Initially people have tried to construct light pipes by cladding the inner wall of a pipe with a normal metal mirror. Due to the albeit small absorption in the metal, which is multiplied by the numerous reflections, these light pipes were very lossy. A nice introduction into these attempts is given by Gambling [16]. Luckily, mirrors can also be manufactured from isolators, materials like e.g. glasses that do not conduct well electricity and hence can have a much smaller light absorption than metals. For a good normal incidence reflection at least two isolators have to be stacked in a periodic way, with a layer thickness of around $\lambda/4n$ with λ the wavelength of the reflected light and n the refractive index of the isolator layer. As the wavelength of visible light is less than a micrometer, constructing these mirrors is a high-tech occupation. A simple stacking of layers, see figure 1.7, leads to a mirror for light rays coming in under a well-defined angle or small range of angles. To extend this angle range, the refractive index difference between the two isolators can be increased, and the stacking can be made two- instead of one-dimensional, see figure 1.7. A logic extension is to construct three-dimensional stacks resembling a NaCl crystal lattice, but with the optical 'atoms' a thousand times bigger than in an molecular lattice.

As these structures make photons behave similar to how electrons behave in an atomic crystal lattice, they are called photonic crystals [17][18]. In particular, just as electrons cannot have certain ranges of

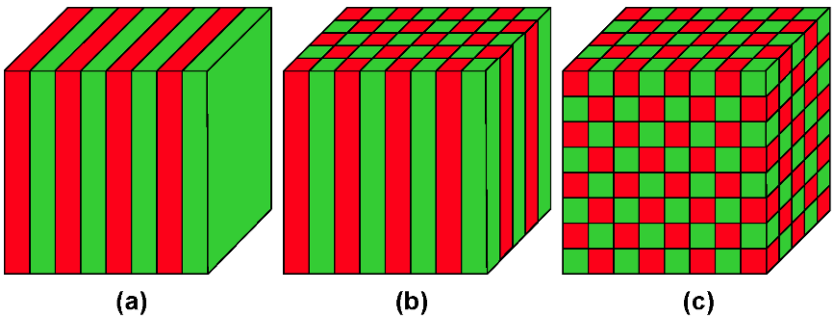


Figure 1.7: Mirrors made out of a **(a)** 1-dimensional **(b)** 2-dimensional and **(c)** 3-dimensional stack of dielectric material.

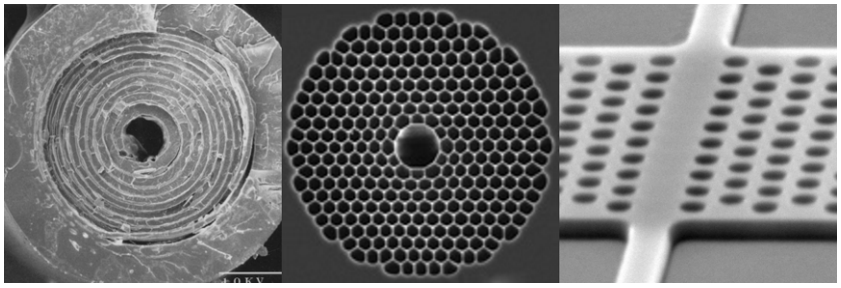


Figure 1.8: Examples of photonic band gap guiding: **Left:** hollow core fiber with stacked dielectric mirror (©Omniguide). **Middle:** hollow core fiber with structured air-silica mirror (©CrystalFiber). **Right:** planar waveguide with two-dimensional photonic crystal as horizontal mirror.

energy within a atomic crystal lattice and for a certain propagation direction, photons also cannot have certain energies within the crystals of figure 1.7 and for certain propagation directions. Analogous to electronics, where the electron energy range is called a band gap, in photonics it is called a photonic band gap (PBG). In contrast to metal mirrors, these PBG mirrors have successfully been applied for guiding light, which we will call *PBG guiding*. Some examples are shown in figure 1.8.

Although one-dimensional and three-dimensional stacks could in theory be placed around a planar waveguide on a photonic chip, in most of the cases a two-dimensional approach as in figure 1.8 (right) is chosen, because it is compatible with standard fabrication technologies.

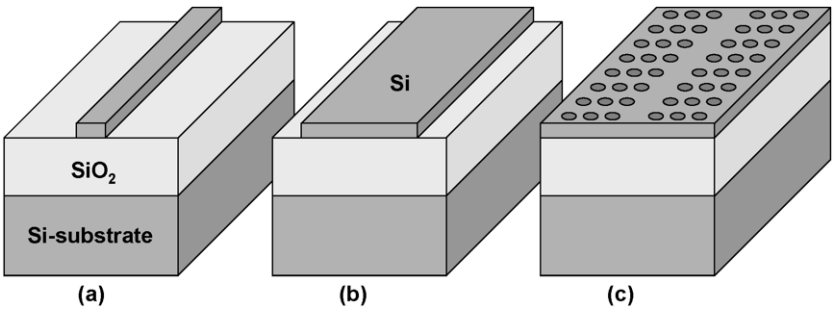


Figure 1.9: Various waveguide types that can interexist within the same PIC. (a) narrow TIR waveguide (b) broad TIR waveguide (c) PhC waveguide.

1.4 Optical Transitions and Coupling

In photonic chips, various waveguide types differing both in geometry and in guiding principle can coexist. A possible example of the coexistence of different waveguide types is the use of photonic crystal waveguides to allow for sharp lossless bends and TIR waveguides for the stretches between these bends. Different waveguide geometries can appear together when a monomodal wire is applied in a bend, but somewhat broader multimodal waveguides, having less loss, in the delay lines or connection lines of a photonic chip.

But also in the connection of an optical chip to the outside world, light spots with different sizes and geometries have to be coupled, like the coupling between an optical fiber and a planar waveguide or between a semiconductor laser diode and a planar waveguide.

As light is to be guided between all these different waveguides, see figure 1.9, the transitions between them need to be designed very carefully to reduce transition losses. The optical component accomplishing this transition is called a taper, a coupler or a spot-size converter.

Sometimes linguistically a difference is made between *a waveguide taper* and a *waveguide coupler*, with the former being the connection between two different waveguides with the same optical axis, and the latter being a component that redistributes light between two spatially separate, parallel waveguides. However, this is not at all a general convention and scientific journals often do not even offer both as available keywords. In what follows, both terms will point to the same component, being the connection of two waveguides with the same optical axis.

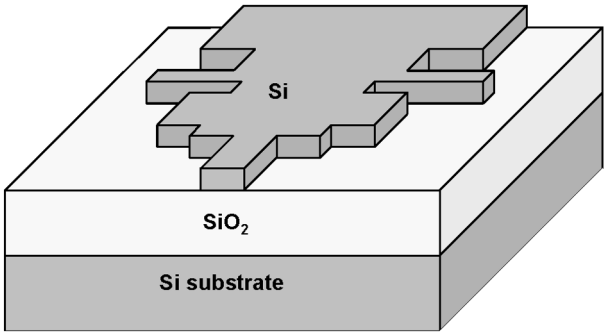


Figure 1.10: A generic taper structure consisting of various waveguide sections between a narrow and a broad waveguide.

Optical tapers can be roughly divided in two groups: three-dimensional (3-D) tapers and planar tapers. Where the former have to change a spot geometry in two dimensions, like transforming a circular mode into a rectangular one, the latter only changes the spot size in one direction, e.g. coupling between the narrow and the broad waveguide in figure 1.9. Although not a universal convention, we will reserve the term spot-size converter for the 3-D case and use the terms coupler and taper interchangeably for the planar case.

1.5 Purpose and Outline of this Work

In this work we will focus on the study of more compact planar waveguide tapers. To couple between the narrow and the broad TIR waveguides of figure 1.9 nowadays gradually opening waveguide tapers are in use with linear or parabolic shapes. Instead an ab-initio optimization of the taper structure in figure 1.10 is to be performed. Good parameters for the waveguide sections that are placed between the narrow and broad waveguides will be determined that result in a good coupling.

After determining how a well-performing coupler can be designed, the same algorithms will be applied for the coupling between waveguides with a different guiding principle, more specifically, between a TIR waveguide and a photonic crystal waveguide.

The decision to limit this study to planar tapers should not be seen as a restriction. It will be explained how a compact planar taper in combination with a vertical fiber coupler can together form a compact spot-size converter.

The text of this book is structured in the following way. In chapter 2 the history and the existing literature on general tapers is briefly discussed. Some special attention is reserved for the more recent research, of which this study is a logical continuation.

Chapter 3 gives a brief overview of the optical simulation tools and the mathematical optimization tools that will be applied further on. The actual optimization and the resulting taper designs are discussed in detail in chapter 4 and compared with existing designs. Three different methods to compose a taper will be tested and compared to each other. Afterwards one of the methods is selected to be tested in the design of a photonic crystal taper.

Besides knowing how to design an efficient taper, understanding how and why it works is also important. The in-depth study of the optical mechanisms underlying the taper working is the subject of chapter 5.

A short introduction in the realization of the designs in silicon-on-insulator is given in chapter 6. The measurement set-up is briefly discussed and the best obtained measurements and the applied data processing make up the rest of the chapter.

In the last chapter the general conclusions of this work are repeated and possible future extensions are proposed.

1.6 Publications

The results obtained within this work have been published in various papers and presented at various conferences. This paragraph gives an overview of the publications. They can also be found in the publications section of the photonics research group website : `http://photonics.intec.ugent.be`. These publications may be downloaded for personal use only. For any other use, the permission of the copyright owner (usually the publisher of the journal) must be obtained.

The following papers have been published in international peer reviewed journals or are accepted for publication:

1. B. Luyssaert, P. Bienstman, P. Vandersteegen, P. Dumon, R. Baets, "Efficient Non-Adiabatic Planar Waveguide Tapers," *Journal of Lightwave Technology*, vol. 23, pp. 2462-2468, August 2005.

2. B. Luyssaert, P. Vandersteegen, D. Taillaert, P. Dumon, W. Bogaerts, P. Bienstman, D. Van Thourhout, V. Wiaux, S. Beckx, R. Baets, "A Compact Horizontal Spot-Size Converter Realized in Silicon-on-Insulator," *IEEE Photonics Technology Letters*, vol. 17, pp. 73-75, January 2005.

3. P. Sanchis, J. Martí, B. Luyssaert, P. Dumon, P. Bienstman, R. Baets, "Analysis and design of efficient coupling in photonic crystal circuits," *Optics and Quantum Electronics*, vo.. 37, pp. 133-147, January 2005.

4. W. Bogaerts, R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, D. Van Thourhout, "Nanophotonic Waveguides in Silicon-on-Insulator Fabricated with CMOS Technology," *Journal of Lightwave Technology*, vol. 23, pp. 401-412, January 2005.

5. W. Bogaerts, P. Dumon, D. Taillaert, V. Wiaux, S. Beckx, B. Luyssaert, J. Van Campenhout, D. Van Thourhout, R. Baets, "SOI Nanophotonic Waveguide Structures Fabricated with Deep UV Lithography," *Photonics and Nanostructures: Fundamentals and Applications*, vol. 2, pp. 81-86, August 2004.

6. P. Sanchis, P. Bienstman, B. Luyssaert, R. Baets, J. Martí, "Analysis of butt coupling in photonic crystals," *Journal of Quantum Electronics*, vol. 40, pp. 541-550, May 2004.

7. P. Dumon, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. Van Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. Van Thourhout, R. Baets, "Low-loss SOI photonic wires and ring resonators fabricated with deep UV lithography," *IEEE Photonics Technology Letters*, vol. 16, pp. 1328–1330, May 2004.

8. W. Bogaerts, D. Taillaert, B. Luyssaert, P. Dumon, J. Van Campenhout, P. Bienstman, D. Van Thourhout, R. Baets, V. Wiaux, S. Beckx, "Basic structures for photonic integrated circuits in silicon-on-insulator," *Optics Express*, vol. 12, pp. 1583–1591, April 2004.

9. W. Bogaerts, V. Wiaux, D. Taillaert, S. Beckx, B. Luyssaert, P. Bienstman, R. Baets, "Fabrication of Photonic Crystals in Silicon-on-Insulator Using 248-nm Deep UV Lithography," *IEEE Journal on*

Selected Topics in Quantum Electronics, vol. 8, pp. 928–934, August 2002.

The following papers have been presented at international conferences and are published in the conference proceedings:

10. B. Luyssaert, P. Vandersteegen, W. Bogaerts, P. Dumon, P. Sanchis, J. Martí, R. Baets, "A Versatile Optical Spot-Size Converter Design," presented at the European Conference on Optical Communication (ECOC), Stockholm, Sweden, pp. 468-469, September 2004.

11. B. Luyssaert, P. Vandersteegen, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, R. Baets, "Compact Photonic Spot-Size Converter," in proceedings of Photonic and Electromagnetic Crystal Structures (PECS-V), Japan, pp. 193, March 2004.

12. B. Luyssaert, R. Baets, "2-D simulations of various compact spot-size converters," in proceedings of Photonic and Electromagnetic Crystal Structures (PECS-IV), Los Angeles, United States, pp. 62, 2002.

13. W. Bogaerts, P. Dumon, J. Wouters, P. Jaenen, S. Beckx, V. Wiaux, D. Van Thourhout, D. Taillaert, B. Luyssaert, R. Baets, "Silicon-on-Insulator Nanophotonics," in proceedings of SPIE Optics and Optoelectronics, Poland, (to be published).

14. R. Baets, P. Dumon, W. Bogaerts, D. Van Thourhout, B. Luyssaert, D. Taillaert, P. Bienstman, V. Wiaux, S. Beckx, J. Wouters, "Silicon-on-insulator platform for WDM-components," in proceedings of the 2004 IEEE LEOS Annual Meeting, Puerto Rico, pp. 741-742, 2004.

15. R. Baets, W. Bogaerts, P. Dumon, G. Roelkens, I. Christiaens, K. De Mesel, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, D. Van Thourhout, Wiaux, Vincent, Wouters, Johan, Beckx, Stefan, "Integration of photonic functions in and with Silicon," ESSDERC/ESSCIRC, Belgium, 2004.

16. P. Bienstman, W. Bogaerts, P. Dumon, D. Taillaert, B. Luyssaert, R. Baets, "Nanophotonics based on high-index-contrast photonics devices," OECC/COIN, Japan, pp. 14E4-1, 2004.

17. P. Sanchis, J. Martí, P. Bienstman, B. Luyssaert, R. Baets, "Analytic expressions for transmission and reflection from semi-infinite photonic crystal waveguides," in proceedings of Photonic and Electromagnetic Crystal Structures (PECS-V), Tu-P31, Japan, 2004.

18. P. Sanchis, J. Martí, B. Luyssaert, P. Dumon, P. Bienstman, R. Baets, "Analysis and design of efficient coupling into photonic crystal circuits," OWTNM04, Ghent, Belgium, pp. 27, 2004.

19. W. Bogaerts, P. Dumon, J. Van Campenhout, V. Wiaux, J. Wouters, S. Beckx, D. Taillaert, B.F.J. Luyssaert, D. Van Thourhout, R. Baets, "Deep UV Lithography for Planar Photonic Crystal Structures," in proceedings of the IEEE LEOS Annual Meeting 2003, United States, p.754-755, 2003.

20. W. Bogaerts, V. Wiaux, P. Dumon, D. Taillaert, J. Wouters, S. Beckx, J. Van Campenhout, B. Luyssaert, D. Van Thourhout, R. Baets, "Large-scale production techniques for photonic nanostructures," in proceedings of SPIE, 5225, United States, pp. 101-112, 2003.

21. W. Bogaerts, P. Dumon, J. Van Campenhout, V. Wiaux, J. Wouters, S. Beckx, D. Taillaert, B. Luyssaert, D. Van Thourhout, R. Baets, "Deep UV Lithography for Planar Photonic Crystal Structures," in proceedings of the IEEE LEOS Annual Meeting 2003 (invited), Tucson, USA, pp. 754–755, 2003.

22. R. Baets, P. Bienstman, W. Bogaerts, D. Taillaert, B.F.J. Luyssaert, "Coupling issues in strongly confined nanophotonic waveguides," Digest of the LEOS Summer Topical Meetings, Canada, pp. 41-42, 2003.

23. R. Baets, P. Bienstman, W. Bogaerts, D. Taillaert, B. Luyssaert, "Coupling issues in strongly confined nanophotonic waveguides," ICTON 2003, Poland, pp. 191-192, 2003.

The following papers have been presented at summer schools or local symposia:

24. B. Luyssaert, P. Bienstman, R. Baets, "A versatile spot-size converter design," IEEE/LEOS Benelux Annual Symposium 2004, Ghent, Belgium, pp. 99-102, December 2004.

25. B. Luyssaert, R. Baets, "A compact spot-size converter," presented at The Rank Prize Fund Mini-Symposium on Passive Optical Components, Grasmere, UK, pp. 16, June 2003.

26. B. Luyssaert, P. Bienstman, R. Baets, "2D simulations of various compact spot-size converters," IEEE/LEOS Benelux Annual Symposium 2002, Amsterdam, The Netherlands, pp. 203, December 2002.

27. B. Luyssaert, R. Baets, "2D simulations of various compact spot-size converters," Workshop on 2D photonic crystals, Monte Verità, Ascona, Switzerland, pp. I-09, August 2002.

28. B. Luyssaert, R. Baets, "2-D simulations of various compact spot-size converters," presented at the IEEE/LEOS Benelux Photonic Crystal Workshop, Ghent, Belgium, May 2002.

29. B. Luyssaert, D. Taillaert, P. Bienstman, R. Baets, "Compact mode converters for PICs," IEEE/LEOS Benelux Annual Symposium 2001, Brussels, Belgium, P56, December, 2001.

30. B. Luyssaert, R. Baets, "2D simulations of various compact spot-size converters," NKT Summer School in Advanced Photonics, Copenhagen, Denmark, pp. I-09, August 2001.

Chapter 2

Optical Tapers

In this chapter, the history of the optical taper will be shortly outlined using the available scientific literature. As 2-D and 3-D tapers have not always been treated separately, the distinction is only made when necessary.

2.1 General Optical Taper Literature

The research on optical waveguides gets a real stimulus with the publication of the paper by Kao and Hockham [19] in 1966. In it they discuss which dimensions and refractive indices the core and cladding of a "dielectric-fiber surface wave-guide for optical frequencies", now an optical fiber, preferably has and they claim that an attenuation of $20\ dB/km$ would be achievable by removing impurities from the glass. This paper convinces a lot of research groups to start working on this promising technology, which in combination with a laser, demonstrated just a few years earlier, can possibly form the highly wanted long-distance high-bandwidth solution the entire telephony business was then looking for. At Bell Labs, Dietrich Marcuse writes a number of theoretical papers starting in '69 on the effect of a non-planar (with respect to the optical axis) core-cladding boundary.

A tapered optical waveguide is first mentioned in [20], where starting from the radiation of a small abrupt width step, the radiation loss of a profiled waveguide is derived. The starting waveguide has to be monomodal (for the chosen polarization, as slab waveguides are described) and has to taper to smaller widths. For this specific case, the

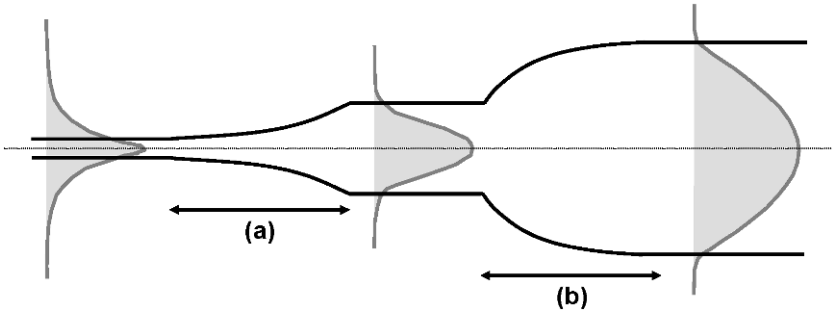


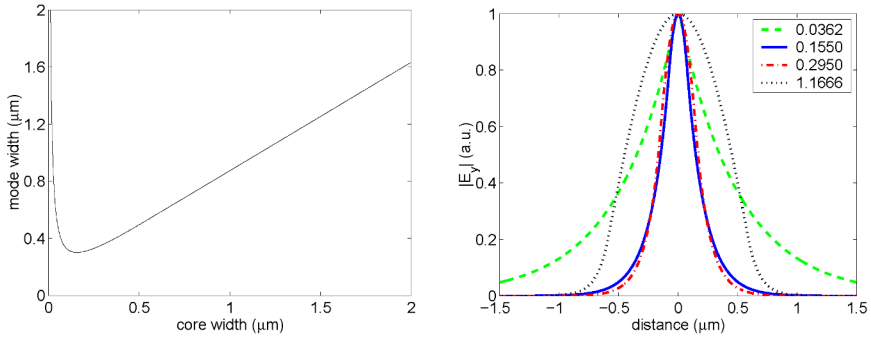
Figure 2.1: Different taper types: **(a)** an inverted taper to couple between a badly-confined and a well-confined monomodal waveguide and **(b)** a regular taper to couple between a monomodal and a multimodal waveguide

author concludes that "a linear taper radiates more on its narrower portion where the field is less tightly guided" and proposes to taper more strongly in the broader section "to distribute the radiation loss uniformly". This kind of taper will be called *inverted taper* in what follows, see figure 2.1. It is important to note that down-tapering a waveguide from a width where the mode is well-confined and monomodal to narrower widths can both increase and decrease the mode width, depending on how narrow the waveguide becomes. Figure 2.2 illustrates this, by showing the mode width as a function of core width and some corresponding mode profiles. For a very narrow core width, the mode will mainly exist outside of the core and be more exponential than gaussian. Although this first paper on tapered waveguides already concludes that a linear taper is far from optimal, the big majority of actually realized tapers are still linear, because they are more easily manufacturable.

A second important paper is the publication in 1975 by Winn and Harris [21], where the coupling from a multimode to a single-mode waveguide using horn-shaped structures is described using eigenmode expansion and coupled-mode theory. Also here the conclusion is that "coupling efficiency can be improved by shaping the walls of the coupler" and the following function is proposed:

$$w(z) = w_0 + 2z \cdot \tan(\theta/2) + A\,(1 - \exp(-z/L)) \tag{2.1}$$

with w_0 the input waveguide width, L the length and θ the full angle of the taper. The first two terms of the equation's RHS describe a linear taper, the third term broadens the taper dominantly at its narrowest



(a) mode width (b) mode profiles

Figure 2.2: Mode width as a function of core width and the corresponding mode profiles for 4 different widths.
n_{core}=2.83, $n_{cladding}$=1.0, λ=1.55 µm.

side. It is numerically verified that this taper form improves the coupling efficiency in respect with a linear taper. This kind of taper will be called *regular taper* in what follows, see again figure 2.1. Quickly this work is continued in [22] where tapered waveguides are applied to relax the alignment tolerances between two coupled waveguides.

The parabolic horn shape is first proposed in [23] as an alternative to the taper shape of formula 2.1 in the form:

$$w(z) = \sqrt{\frac{2\alpha\lambda_0}{n_{core}}z + w_0^2} \tag{2.2}$$

with α a constant less than or equal to unity, and sometimes called the adiabaticity parameter. As the theory in [23] is based on a ray model, the authors have refined it in [24] with the taper form:

$$w(z) = \sqrt{2\alpha\lambda_g z + w_0^2} \tag{2.3}$$

with $\lambda_g = \lambda/n_{eff}$ the material wavelength of the lowest order local waveguide mode. This time the authors apply eigenmode expansion and coupled mode theory. To strictly obtain a parabola, λ_g must be place-independent, so one has to assume that both ground and first even mode are far away from cut-off, making formula 2.3 only valid for coupling between broad multimode waveguides with a strong confinement. The authors note that "in the general case [...] the horn shape will

not be strictly parabolic" and "it is not clear that our parabolic shape has any distinct advantage over a carefully chosen exponential taper such as the one discussed in [21]". In appendix A, starting separately from a ray model, from Gaussian beams and from mode interference, formula 2.3 is derived and it is proven that this parabola is symmetric around the propagation axis.

To our knowledge the 1977 paper [24] by these authors is the first to use the word "adiabatic"[1] in a optical taper context. Coming from thermodynamics and later quantum mechanics, in the latter case the word describes *a system changing so slowly that all the energy that is initially in state* a *will very likely go over to the corresponding state* b *without making any transition* [25]. In optical waveguide theory, this translates to the situation where all the energy is within a certain mode of the waveguide, mostly the ground mode, in every cross-section of the taper. In thermodynamics a process is adiabatic if there is no heat gained or lost. This could translate to optical waveguide theory as the situation where there is no intensity loss, regardless whether inside the taper energy might be redistributed between the higher-order modes. In what follows we will use the word adiabatic only in its strictest quantum mechanical sense.

Going beyond the two-mode approximation, the paper by Lim [26] studies how mode i injected in the narrow part of a linear taper couples to every other mode j and concludes that:

- coupling is highest for $j = i \pm 1$ and $j = i \pm 2$
- coupling is higher for $j > i$
- the steeper the taper, the higher the coupling
- the two-mode approximation is not sufficient for very short tapers
- a reasonably large step size can be used in numerical calculations

with the step size being half the width difference between two consecutive linear waveguide sections.

The above papers all employ eigenmode expansion, mode coupling theory and a staircase approximation to simulate waveguide ta-

[1] coming from the Greek adiabatos, meaning impassable ($\alpha + \delta\iota\alpha\beta\alpha\iota\nu\epsilon\iota\nu$ = not + to go through)

pers. Other theoretical models are applied in Tzoar [27] and Lee *et al.* [28].

In the work of Tzoar [27], instead of a staircase approximation the eiconal or local-normal-mode approximation method is applied to handle longitudinally dependent coefficients. Using the approximation of very gradual tapers, the paper reaches the same conclusion as the step-approximation models. More information on the local normal modes can be found in [29].

The publication by Lee *et al.* [28] proposes a equivalent-waveguide concept to design the completely adiabatic tapered waveguide and builds on some earlier papers that apply refractive index tapering with or without dimension tapering. If both index and dimensional tapering can be used without restrictions (actually meaning that within every point of a structure a different refractive index can be set) this paper outlines how the complete adiabatic taper has to be calculated. The authors use a series of theoretical conformal mappings, better known from bend waveguide theory, to transform a single-mode straight waveguide into an equivalent tapered configuration. As the method inherently introduces curved-phase fronts within the taper structure, which indicates higher-order mode excitation, the word adiabatic is used in its broad thermodynamical sense. The theory of this paper is reformulated by Ash *et al.* [30] and a simple example of such a conformal mapping is given and computed. Again it is concluded that the required refractive index distribution is purely theoretical and out of reach of any existing fabrication process.

Instead of looking for the best form to taper down on an as short as possible length, some authors have investigated tapers in the limit of very long taper lengths. In [29] and more explicit in [31] it is proven that if the taper length becomes infinitely long the coupling coefficients between the mode under study and every other mode become zero, provided that this studied mode is guided, non-degenerate and lossless. This also means of course that for very long tapers, the taper form itself is not important anymore. As long as in every cross-section of the taper the above conditions for the mode are met, the taper will be lossless or adiabatic.

This section can be summarized in the following way. When tapering from a monomodal waveguide to an even smaller waveguide, the tapering can go faster in the broad section. In the other case, when tapering from a monomodal or a multimodal to a multimodal waveguide, the opposite is true, so tapering can start of rather fast but has to

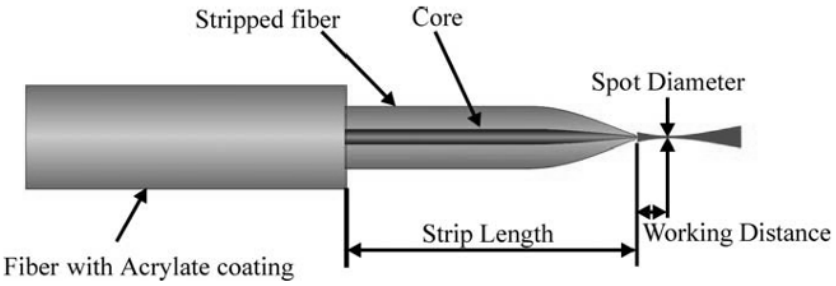


Figure 2.3: Drawing of a lensed fiber ©HTSKorea.

smoothen towards the broader waveguide. An analytic expression for the optimal taper form can only be given in a very specific case (very broad highly multimodal waveguides with good confinement) where it is a parabola. In the more general case, an analytic expression does not exist and simulations are unavoidable.

2.2 3-dimensional Tapers or Spot-Size Converters

2.2.1 Lensed Fiber

Although the papers of the previous section mostly use slab waveguide theory, the first experiments on tapered waveguides were done with three-dimensional waveguides, like Ti-diffused $LiNbO_3$ waveguides [32], laser-written waveguides in glass [33] or Ag-diffused soda-lime-glass waveguides [34].

As the first optical links only consist of a source, a fiber and a detector, much of the early research is on the laser-fiber coupling. Because detectors are generally larger than the fiber mode, diffraction and an accurate alignment do the job at the detector side of the link. At the laser-fiber interface one can choose whether to adapt the laser or the fiber or use micro-optics like micro-lenses for a good coupling between the broad circular fiber mode and the much smaller asymmetric laser mode. The use of micro-optics adds additional expensive alignment steps to the package, and is therefore not preferred. Adapting the fiber is frequently done in the form of tapered (or lensed) fiber, see figure 2.3, and is a practical way to give the laser and the fiber mode the same size. It has however three disadvantages: the waveguide mode remains asymmetric and will not couple completely to the circular fiber mode,

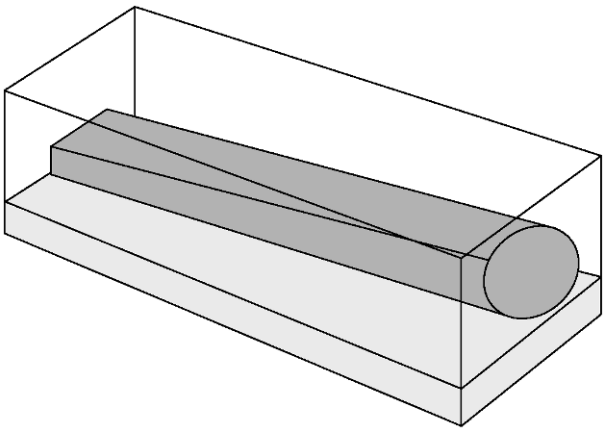


Figure 2.4: Conceptual taper to couple between a edge-emitting laser and a fiber. The real part of the refractive index is according to the gray scale. The figure is not to scale, waveguide mode size is in the order of microns, while taper lengths in the order of millimeters.

a piece of fiber needs to be incorporated in the laser package and the alignment tolerance is very low.

2.2.2 Integrated Spot-Size Converters

In principle, designing a three-dimensional taper between a laser and a fiber is not that hard. Building on the adiabatic theorem, a simple structure like in figure 2.4 suffices, if long enough. However, as the techniques to fabricate optical components are mostly planar, the structure in this figure is not a realistic option. Several different structures have been proposed to manufacture three-dimensional tapers using well-established techniques. More than 15 different taper structures that can be monolithically integrated with III-V semiconductor devices are described in an overview paper by Moerman *et al.* [35]. As these structures need not only to change the size of the asymmetric laser mode both also the mode shape, they are often called spot-size converters, or shortly SSC.

Some of the spot-size converters need rather complex fabrication like local (re)growth, dynamic or shadow etching and are therefore less suited for mass fabrication. A nice example of such an integrated spot-size converter, although needing regrowth, has been studied by Willems in our research group [36]. To expand a waveguide mode in

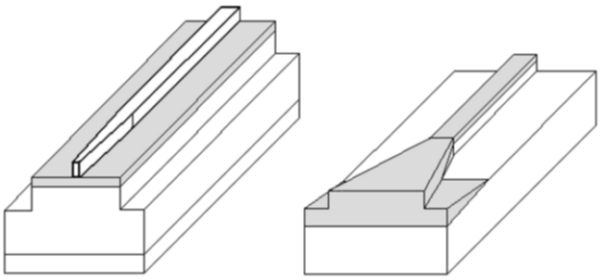


Figure 2.5: Two typical three-dimensional integrated tapers. **Left** with an inverted taper. **Right** with a regular taper.

the vertical direction, it couples to other core layers, one above and one below. By choosing the length of the multi-mode section adequately this results in a broad mode excitation and better coupling to a fiber. A lot of work is still going on to fabricate monolithically integrated three-dimensional tapers with as few extra fabrication steps as possible, compared to standard edge-emitting lasers. In our group the recent work by Galarza [37] in InP and De Mesel [38] in GaAs is in this context.

All these different taper proposals can be roughly categorized into two main types, as illustrated in figure 2.5. Both have at one end a waveguide supporting the laser mode and the other end a broad waveguide with a nearly circular fiber-matched mode, but differ in how these two waveguides are coupled. The left structure uses an inverted taper to bring the laser waveguide down to a narrow tip. Towards this tip, the mode inflates, exists mainly outside the waveguide and starts feeling the broad waveguide, to which it is drawn. At the end all the energy is in the broad waveguide mode. In the right structure, a regular taper couples the two waveguides, which have different dimensions both horizontally and vertically, requiring a taper in both directions. Both designs have their drawbacks, in the first a narrow tip has to be defined, requiring high resolution techniques like e-beam lithography, and in the second the tapering in the vertical direction is a problem for planar techniques.

Although it fits into the above inverted taper category, the spot-size converter recently proposed by NTT [39], deserves to be mentioned separately due to its very low conversion loss and its compatibility with mass fabrication, see figure 2.6. Building upon earlier research by NTT [40] where this kind of taper has been integrated into an edge-emitting 1.3-μm laser diode, the more recent work is aimed towards the coup-

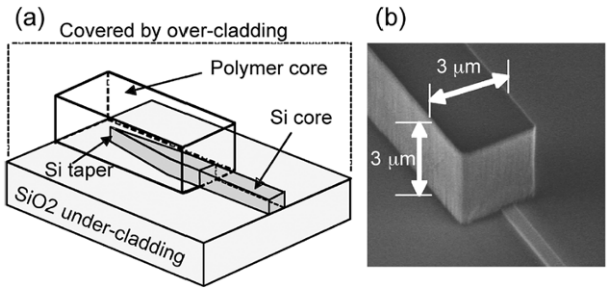


Figure 2.6: **(a)** Schematic of NTT spot-size converter. **(b)** Scanning electron micrograph of the junction point of the Si wire and the polymer core. ©JSTQE

ling of an optical fiber with passive silicon-on-insulator micro-photonic circuitry. A narrow 0.4 µm silicon wire is linearly tapered down to a tip of just 80 nm that is covered with a square waveguide with dimensions of 3x3 µm and a refractive index of 1.5. This square waveguide is covered with an over-cladding having a refractive index of around 1.45. In the first version of this spot-size converter a medium and high refractive index polymer were used for these two materials [41]. In the latest version, inorganic materials like silica (SiO_2) and siliconoxynitride (SiON) replaced these because of their better absorption (no water peak, see section 1.2) and humidity resistance properties [39]. A total conversion loss of only 0.5 dB per connection has been achieved, with one connection being the fiber - square waveguide - silicon wire transition. One drawback of this spot-size converter is that the square broad waveguide is designed for good coupling to specialty fiber, namely of small-core fiber with a mode size of 4.3 μm. The coupling loss for a conventional fiber with a mode-field diameter of 9 μm was found to be about 2.5 dB per connection for the inorganic converter. Recently a new design for this taper type has been proposed having a wider taper tip, manufacturable using deep UV lithography [42].

While in the more recent work the inverted taper is mostly linear, in the older work non-linear tapers of the form

$$w(z) = w_0 + (w_L - w_0)(1 - z/L)^n \tag{2.4}$$

are used, with the best results for $n = 2$ or $n = 3$. Due to the very low absorption of the inorganic compounds. However, the taper size itself does not strongly affect the overall component loss and linear tapers, which are easier to fabricate, can now be chosen over a shorter

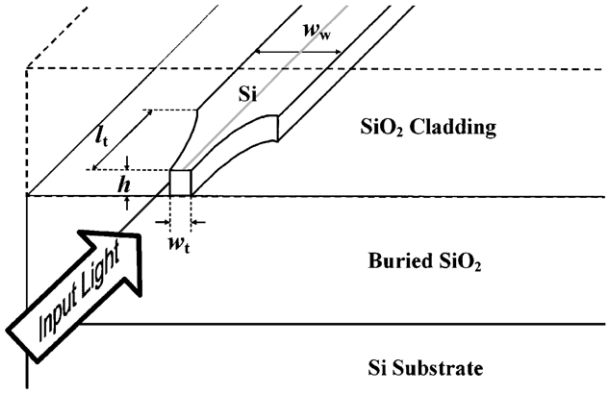


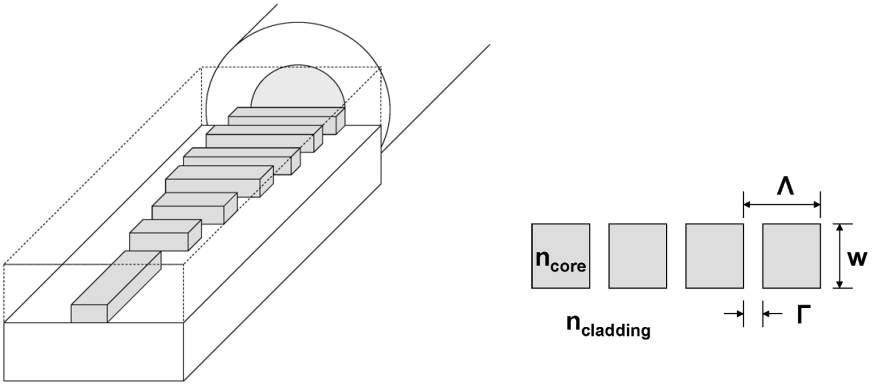
Figure 2.7: Schematics of an inverted lateral taper for direct coupling to an optical fiber. ©OPTICS LETTERS

component length.

Instead of defining an intermediate square waveguide, lateral tapering down to a narrow tip, covered by an over-cladding alone, can achieve good coupling to fiber by itself, see figure 2.7. This taper type was first proposed and demonstrated by NTT in 1993 [43] with an InGaAsP core and an InP cladding and substrate. Because the tapering is only lateral the fabrication is compatible with planar manufacturing techniques. The waveguide is shown to have a total insertion loss of 2.6 dB and a large $\pm$2.5 µm (1-dB) misalignment tolerance in lateral and vertical direction. The width of the tip, however, is very critical to achieve a good coupling.

In a follow-up paper in 1994 [44], this taper type is described more theoretically. Using the same material system, the authors show that a coupling loss of only 0.5 dB to conventional single-mode fiber can be achieved. To reduce the total length of the component, taper forms as in formula 2.4 are studied, with the best results for $n = 2$ and $n = 3$.

The same taper type but in a different material system, namely silicon-on-insulator, is studied by Almeida *et al.* [45] in 2003. As the waveguide mode is much smaller in this system, the inverted taper can be much shorter too, just 40 μm instead of 700 μm in [44]. This time not only the coupling of the TE-mode but both polarizations are studied. The required tip width for good coupling is very different for both polarizations. While the optimum for TE is around 120 nm, for TM this is just 50 nm, and these optima are also rather narrow, with the 1 dB



(a) Waveguide mode expander

(b) Periodically segmented waveguide

Figure 2.8: Schematics of a periodically segmented waveguide mode expander as proposed in [47] and some parameter definitions of a segmented waveguide.

tolerance interval just 15-20 nm at each side of the optimum. The optimal coupling for a TE-mode between the tip and a small-core fiber (mode diameter of 4.6 µm) gives a (theoretical) coupling loss of just 0.26 dB, which is even better than in the InP case because the effective indices of the tip mode and the fiber mode are closer together. This structure could be adapted for coupling to conventional fiber, but this would bring an even narrower tip, less tip width tolerance and thicker cladding layers, making the fabrication harder.

2.2.3 Waveguide Mode Expander

Another remarkable taper type that exploits only planar techniques but achieves three-dimensional mode conversion is the periodically segmented waveguide mode expander as proposed by Weissman [46], see figure 2.8(a).

A segmented waveguide has a core that is sometimes interrupted by sections that have the same refractive index as the (over-)cladding. In a theoretical paper that describes these waveguides and the corresponding modes in a general way [48], the authors prove that weakly guiding periodically segmented waveguides (PSW) with

$$n_{core} = n_{cladding} + \Delta n \tag{2.5}$$

have the following properties:

- Δn_{eff} of a PSW is equal to that of a continuous waveguide with an averaged index difference $\Delta n' = \Delta n\,(1 - (\Gamma/\Lambda))$, meaning that the effective-index difference is independent of the period Γ and is roughly linear with the segmentation ratio $(1 - \Gamma/\Lambda)$.

- Reflection bands are numerous in the (Γ, λ) plane. But, as the fractional bandwidth is only a few percent, there is enough design flexibility to avoid reflection.

- Losses of the ground mode increase monotonically with Δn and oscillate with w, Γ, λ and $(1 - (\Gamma/\Lambda))$.

The first taper proposed by the authors [46] consists of a PSW with a fixed core width, fixed period but a segmentation ratio that changes between 1 and the value that generates the mode with the smallest mismatch with the optical fiber mode. Using the beam propagation method it is proven that this taper effectively broadens the waveguide mode in both dimensions without considerable loss. The more complicated taper of figure 2.8(a) is a combination of a PSW with a fixed width and a PSW with a linear taper envelope. Both sections have a fixed period, the straight section has a fixed segmentation ratio and the tapered section has a segmentation ratio $(1 - \Gamma/\Lambda)$ that changes linearly between that of the straight section and 1.

The design is numerically verified using BPM [47] for a square silica 3x3 µm-waveguide with Δn=0.015 and n_0=1.48 and yields for a taper with a total length of 1 mm an total conversion loss of 0.38 dB. For an InP strip 0.3x1.5 µm-waveguide with Δn=0.085 and n_0=3.17 the coupling loss is reduced from 4 dB to about 0.7 dB using a taper with a total length of 600 µm. As the ground mode propagation loss of a PSW increases with increasing Δn, this kind a taper is not practical anymore for a material system with very high refractive index contrast.

2.2.4 Optimized Waveguide Mode Expander

In [47] the author mentions the possibility of also changing the segmentation period Λ along the taper, which would strictly make it non-periodic. This extra parameter makes the design of the taper much harder.

Spühler *et al.* [49] have worked out this idea, but instead of taking two sections that are optimized separately, they have optimized the entire coupler at once, see figure 2.9. To avoid putting restrictions on

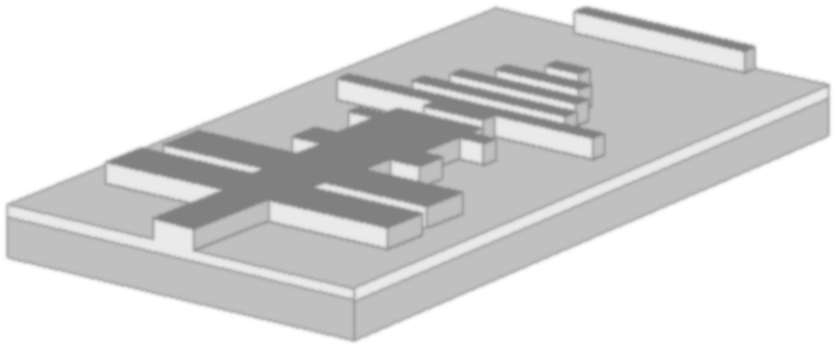


Figure 2.9: Schematics of the optimized non-periodic segmented waveguide taper as proposed by Spühler *et al.* [49]. The fiber position is at the right. Overcladding has been omitted for clarity.

the design, almost no prior knowledge is applied. A structure is divided into 70 sections with a length (propagation direction) of 3 µm, and a width (transversal horizontal) between 1.0 and 21.0 µm in steps of 0.5 µm, or a negative width, which means that the section is omitted during calculation. These numbers give a total of around $42^{70} \approx 4 \times 10^{113}$ possible taper structures. To find a good design within this huge multi-dimensional parameter space, the authors have applied an evolutionary optimization procedure, in combination with BPM for calculating the structures. The resulting taper, see figure 2.9, looks rather random and concepts like segmentation period or ratio are not applicable anymore, but it improves the coupling efficiency with 2 dB up to a coupling loss of only $-1.3\,dB$.

2.2.5 Optimization and Integrated Optics

With ever more available and cheaper computing power, optimization algorithms are more and more applied in engineering and integrated optics is no exception. Problems that cannot be solved analytically and that are too complex to gain insight, are being optimized using advanced algorithms. One class of these algorithms is the genetic or evolutionary algorithm, which tries to mimic the course of natural evolution mathematically. For an overview of its use in electromagnetics, like complex antenna and filter design, see [50].

More specific in the field of integrated optics, genetic algorithms have been applied to optimize the following components (list is by no means complete):

- an integrated spot-size converter in SiON/SiO_2 [51]
- a non-periodic grating coupler [52]
- a periodic grating coupler in SOI [53]
- a photonic crystal spot-size converter [54]
- a hybrid conventional and photonic crystal waveguide bend [55]

This trend will certainly continue in the future as computing power will become even cheaper and the field of mathematical optimization will probably discover even stronger search algorithms.

2.3 Two-dimensional Tapers

In the following, we will use the term *two-dimensional taper* for a structure that only tapers horizontally and where the vertical mode size does not change significantly, to exclude a component like the one of figure 2.7.

2.3.1 Literature Overview

Since the original papers of section 2.1, little research has been devoted to two-dimensional tapers, probably because in experiments the total size of a component is not that important a factor and therefore linear tapers that are a few times the adiabatic length, see appendix A, are widely used.

Long parabolic tapers have small transmission losses but determining these losses can still be hard, depending on the exact waveguide and taper parameters. The earliest studies used three-layer waveguides, of which the waveguide modes can be calculated analytically. For more complex waveguides, numerical methods are necessary and the beam-propagation method has been applied a lot in the past, see also appendix A. However, for increasing refractive index contrasts and shorter tapers this method becomes less and less correct and other electromagnetic calculation tools become necessary. A detailed study, conducted in our research group, gives more insight into the correctness of various simulation tools for different two-dimensional tapers [56][57].

In replacement of a smooth parabola, Kiseliov [58] studies the best way to approximate a parabola by a piecewise linear taper form. Although this work is originally meant for millimeter waves, it can be

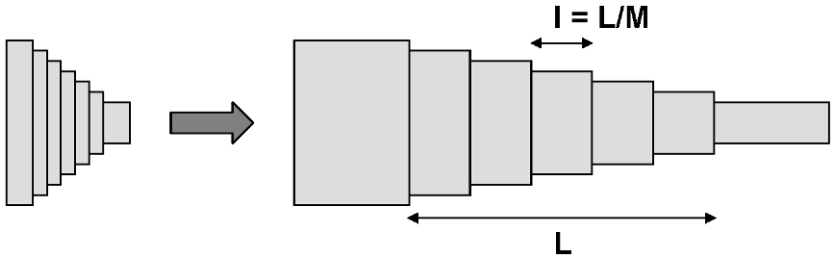


Figure 2.10: Stretching towards long step tapers.

completely translated to optics. For a certain number of linear sections, formulas are derived to determine the coordinates of these sections. Again this theory only applies to waveguides that are highly multimodal and well-confined.

On lithographic masks, a smooth parabola is often approximated by a step taper with small steps. Depending on the resolution of the lithography, this pattern might or might not be transferred onto the real component. Felici *et al.* [59] look at the effect of this possible discretization on the taper losses. For tapers around the adiabatic limit, Lim [26] already concluded that a reasonably large step size may be applied in numerical calculations. If these tapers are stretched to long lengths, while keeping the discretization, as in figure 2.10, additional losses arise, however. More specifically, power loss will blow up for lengths that satisfy the following criterion:

$$L = \frac{2\pi M}{\bar{\beta}_1 - \bar{\beta}_j} \tag{2.6}$$

with M the number of steps, $\bar{\beta}_1$ the average wave number of the ground mode and $\bar{\beta}_j$ the average wave number of the j^{th} mode (forward or backward). The shortest of these resonant lengths is:

$$L_0 = \frac{\pi M}{\bar{\beta}_1} \tag{2.7}$$

In another paper by Felici *et al.* [60] the coupling problem, as in figure 2.11, is studied as an inverse problem. Instead of proposing a certain taper shape and studying how it performs, the authors work the other way around and try to find a refractive index distribution that will couple the ground mode as well as possible between two different waveguides. In contrast with [28] and [30], where the perfect refractive

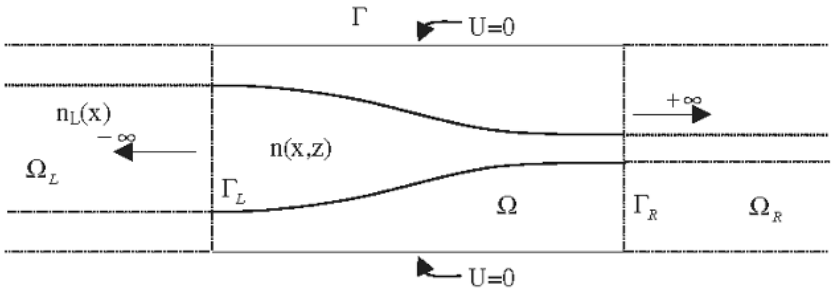
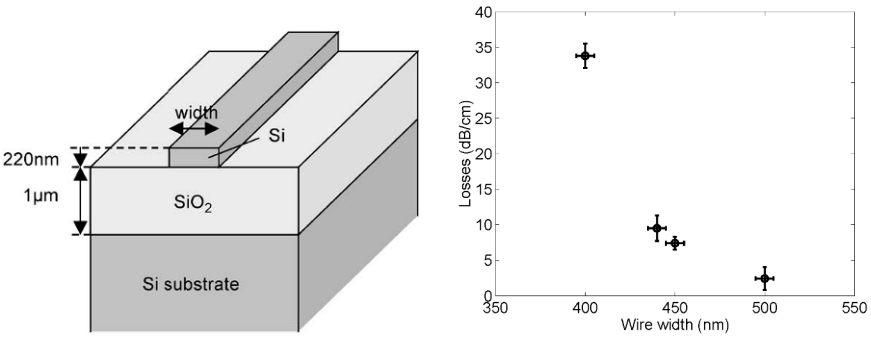


Figure 2.11: The general coupling (inverse) problem as studied in [60].

index distribution problem is calculated regardless whether it can be manufactured or not, this paper studies the remaining optimization after setting certain constraints for $n(x, z)$. As a test case, $n(x, z)$ is chosen as a series of step-index uniform waveguide sections with fixed length, the remaining parameters to optimize are the widths of these sections. Note that this problem is similar to the one studied in [49], see figure 2.9. Important conclusions from the general mathematical analysis of this taper optimization are that:

- an optimal taper does not have to be adiabatic. Inside the taper higher-order modes can be excited, while the entire structure can have a ground-to-ground mode transmission that is almost perfect.
- the optimization is ill-posed, meaning that one can always find an arbitrarily large variation in the optimal solution, which still gives a field arbitrarily close to the optimal field. In the test case this translates to the fact that a narrow waveguide sticking out will not change the field in a significant way.
- this ill-posedness causes numerical instability in the convergence of the optimization, especially if the optimization algorithm uses gradient information.
- numerical stability can be recovered by limiting the search space to a smaller sub-search space. An example is to disallow spikes in the optimization.

Based on the results of this paper, the authors have developed an automatic optimization engine KALLISTOS that is distributed by Photon Design.



(a) silicon-on-insulator wire (b) Wire loss versus wire width

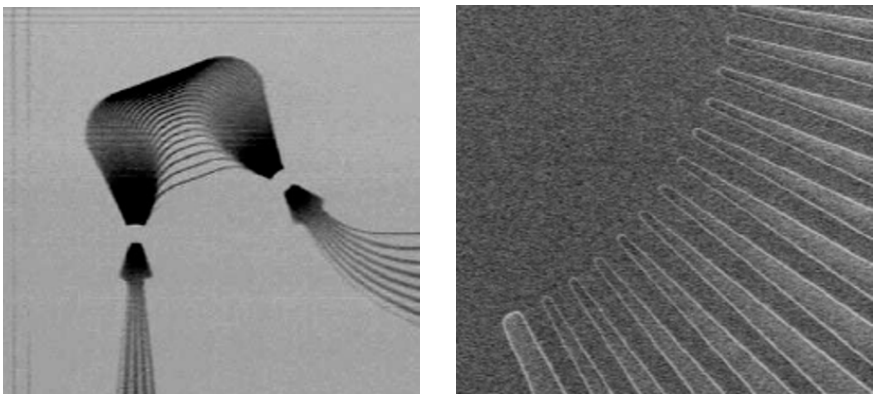
Figure 2.12: Schematic of a silicon-on-insulator wire (©PTL) and its propagation loss as a function of wire width, data taken from [61].

2.3.2 Two-dimensional Taper Applications

While tapering, up until today, has been mainly applied for the coupling of one optical component (like a laser) to an optical fiber, integrated optics will change this picture in the future. An interface to the outside world is still necessary. But with the integration of several optical functions into one chip, two-dimensional tapers also become essential components. This is the case when two different waveguides have to be seamlessly joined.

One possible application is in a complex optical circuit using, for example, SOI wires, see figure 2.12(a). To be able to create a short bend, the wire has to be strictly monomodal, meaning that it supports only 1 guided mode, and as well-confined as possible. This corresponds to a wire width of around 445 nm [62], for which the propagation losses are very high, see figure 2.12(b). Long straight pieces of waveguides, like delay lines, or connections between various components, would benefit from a smaller loss and therefore have to be wider. A taper is then needed to connect the narrow and the somewhat broader waveguide.

Another application of tapers within an optical chip, could be in an arrayed waveguide grating (AWG) or phasar [64], see figure 2.13(a). The transition between the slab region and the array waveguides, see figure 2.13(b), has to be designed very carefully to suppress reflection and to guide as much light as possible into the array waveguides. One technique to smoothen this transition is to apply a double etch process, which has recently been demonstrated in InP [65] but could be ported



(a) Arrayed waveguide grating in SOI (b) Slab region and array waveguides

Figure 2.13: Scanning electron micrographs of an AWG in SOI and a zoom of the transition between the slab region and the array waveguides. Pictures taken from [63].

to other material systems. We believe a compact two-dimensional taper could also be applied in this situation.

Also at the interface between conventional waveguides and photonic crystal waveguides, two-dimensional tapers can be applied. The next section will treat this problem into more detail.

With the recent successful demonstration of a broadband vertical grating coupler [66][67][53], see figure 2.14, that can be integrated in every high-index contrast material system, another important application for two-dimensional tapers has been created. When using cleaved standard optical fiber the waveguide in which the grating (1-D or 2-D) is defined has to be at least as wide as the mode size within the fiber (around 12 µm). The monomodal wires that are preferentially used in these high contrast materials have widths around 0.5 µm. With conventional tapers, the coupling between the wide and the narrow waveguide becomes very long (hundreds of microns) and a compact taper can save a lot of chip area in this case. With the combination of a vertical fiber coupler and a compact planar taper in mind, the coupling between waveguides with widths of 0.5 and 10.0 or 12.0 µm is one of the studied coupling scenarios in the next chapters.

The designs we will explore build on the general conclusions of the theoretical study by Felici [60] and on the optimization algorithm as explored by Spühler [49].

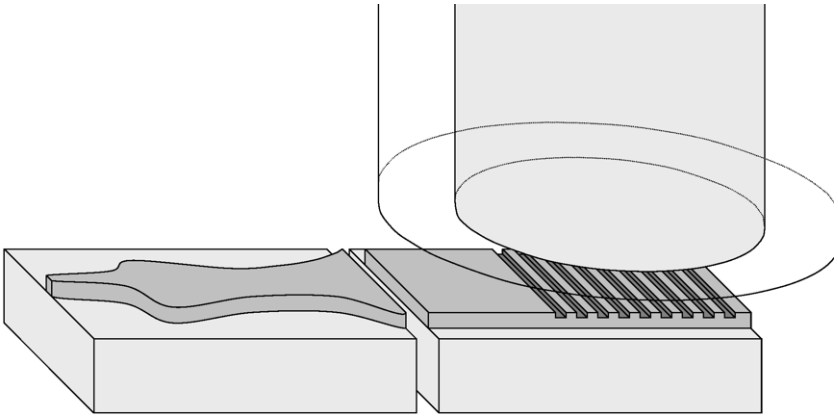


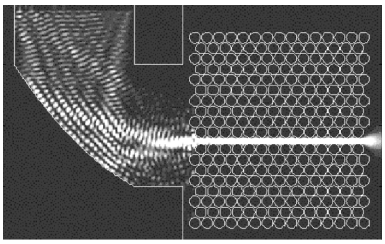
Figure 2.14: Vertical grating coupler between broad waveguide and fiber [66].

2.4 Photonic Crystal Tapers

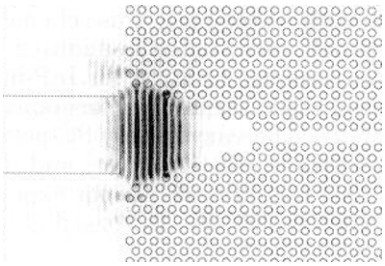
2.4.1 Coupling Into and Out of Photonic Crystal Waveguides

The coupling between an optical fiber and a laser diode could be improved by changing each component, i.e. by using lensed fibers or integrated spot-size converters or both. Similarly the coupling between a conventional waveguide and a two-dimensional photonic crystal (hereafter called PhC) waveguide can be improved by changing the conventional waveguide or the PhC waveguide or both. To indicate the width of a PhC waveguide the conventional notation will be used, i.e. *W1* for a waveguide formed by leaving out 1 row of holes (or pillars) and *W3* when 3 rows are omitted.

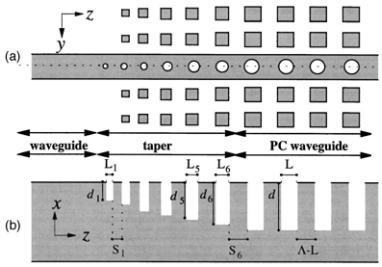
The coupling between a silicon wire and a W1 PhC waveguide is studied using 3-D FDTD by Miyai *et al.* [68]. The difference between air-bridge and silica underclad structures and the influence of the wire width is investigated, as these are the only parameters that easily can be changed when using standard fabrication technology. For wavelengths between 1.50 and 1.55 µm, a coupling efficiency is calculated that exceeds 80% . In Prather *et al.* [69] a PhC W1 waveguide is placed at the focal point of a parabolic mirror, etched into a broad conventional waveguide, see figure 2.15(a). A theoretical efficiency of 90% (2-D FDTD) is reported and first measurements indicate: *light [...] on the output facet* but no quantitative results are given. As this structure is not specific to PhC incoupling it can be seen as a more general two-dimensional taper.



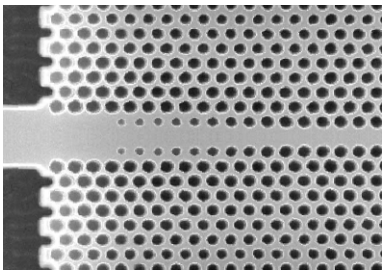
(a) J-coupler [69]



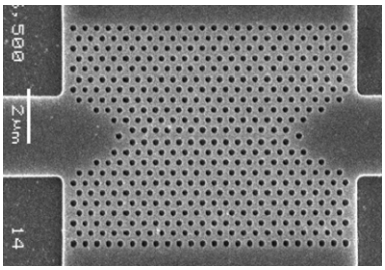
(b) Linear PhC taper [70]



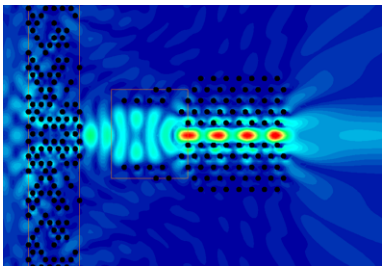
(c) Adiabatic PhC taper [71]



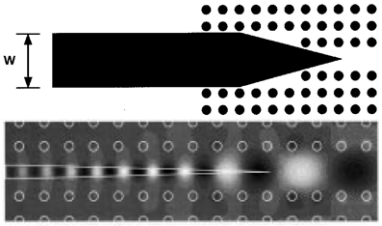
(d) W3→W1 taper [72]



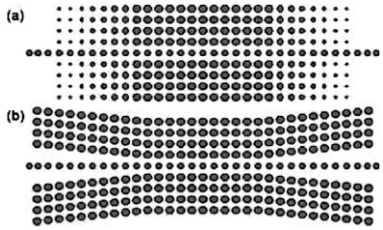
(e) Optimized linear PhC taper [73]



(f) Optimized PhC grating [54]



(g) Adiabatic transition [74]



(h) Bad and good adiabatic taper [31]

Figure 2.15: Various proposals for PhC tapers. Pictures taken from: (a,b,d) ©Opt. Lett., (c,f) ©App. Phys. Lett., (e) ©Phot. Tech. Lett., (g) J. Lightw. Tech. and (h) ©Phys. Rev. E

As it was soon understood that coupling to broad PhC waveguides is easier than to a W1, because of the PhC mode more resembling a conventional waveguide mode, some work has been done on PhC tapers. Linear PhC tapers are studied (2-D FDTD) by Happ *et al.* [70] and are proven to be more efficient than linear tapers of equal length in a conventional waveguide, see figure 2.15(b). The resulting efficiencies are, however, too low to be of any practical use. A first approach to an adiabatic PhC taper [71] gradually changes a wire into a PhC waveguide by simultaneously etching deeper and bigger holes as in figure 2.15(c). This general proposal is continued by a more specific design for a $W3 \rightarrow W1$ taper, figure 2.15(d), by gradually increasing holes, first theoretically [75] and later also as a realized component [72]. Seeking to improve a linear PhC taper, Sanchis *et al.* place one [76] or two [77] additional holes inside the tapered region. Measurements [73] show that these additional defects, see figure 2.15(e), can raise the efficiency from 30% for an unaltered PhC taper to 60%. Instead of really altering the PhC waveguide Witzens *et al.* [78] place a multilayered PhC grating at the interface between a broad conventional waveguide and a slab waveguide. By applying superprism effects and constructive interference between the different PhC rows, they mode match the broad waveguide and the PhC waveguide, which is located inside the slab region. A similar idea is worked out by Sanchis *et al.* [54] but instead of changing the pitch between the different rows of PhC holes, they start from a fixed array of holes. A genetic algorithm then determines whether a hole has to appear or not for the best mode matching. An additional small array of holes focuses the light into the PhC waveguide, see figure 2.15(f).

The structure of figure 2.15(c) [71] is an example of a PhC coupler where both the conventional and the PhC waveguide are changed for a better matching. Other examples are studied by Mekis *et al.* [74] , see figure 2.15(g), and Xu *et al.* [79] and also optimize the coupling between a wire and a PhC made out of high-index pillars.

A very important paper is the one by Johnson [31], where an adiabatic theorem for *grated waveguides* is derived in a very theoretical way. In words, this theorem states (1) the transition has to occur slowly and (2) in every point throughout the taper the propagating mode must be guided and non-evanescent. For every period of the taper the band gap must be calculated and one must check whether the above rule applies. The author uses the term *grated waveguides* as a generalization of photonic crystals, because in the taper region pure periodicity will be lost.

Building on this knowledge he also discusses two transition schemes, see figure 2.15(h), and demonstrates that while the lower one (zipper structure) is truly adiabatic, the upper one is not. This work explains in a theoretical way what had been found earlier in [80].

Although the above list of PhC tapers is probably not complete, it provides a quick overview of the various attempts to create good coupling between a conventional and a PhC waveguide. Other techniques that have appeared in the literature, like the evanescent fiber coupler [81], can also be useful and promising but do not act fundamentally different from similar components that couple to conventional waveguides.

In chapter 4 the coupling between a TIR and a PhC waveguide will be optimized by changing the TIR waveguide side of the interface.

Chapter 3

Methods

In this chapter we will take a closer look at the methods for simulation and optimization that will be applied further on. The eigenmode expansion method is explained and its possibilities to sometimes apply symmetry.

An introduction into genetic algorithms for optimization is followed by the exact implementation of this algorithm used in this work. A special optimization method that we have applied to improve parabolic tapers is also explained in detail.

3.1 Eigenmode Expansion Method

All simulations described in this work have been performed using an eigenmode expansion method (EEM). For a very detailed study of this method, see [82].

To solve Maxwell's equations for a particular structure using eigenmode expansion, first a propagation direction has to be established. As we study two-dimensional tapers, this propagation direction will, of course, coincide with the optical axis of the end waveguides, called the z-axis hereafter. The simulation area has to be finite, which is arranged by enclosing the structure in a rectangular box made out of electric or magnetic walls. These are artificial walls, formed by forcing some of the electric or magnetic field components to be zero. As this also implies perfect reflection at these walls, special boundary conditions have to be set to prevent these reflections. The absorbing boundary condition most widely used today, in a lot of electromagnetic simulation methods, is a perfectly matched layer or PML, again [82]. The combi-

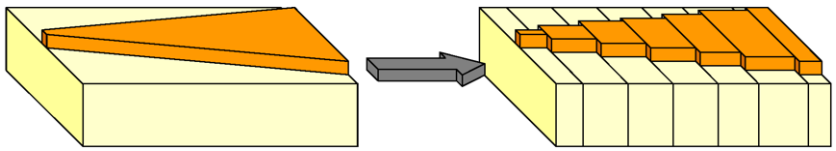


Figure 3.1: Discretization of a structure in slabs that are invariant in the z-direction.

nation of electric or magnetic walls and PML offers the advantage of a finite simulation volume while, from the inside of the box, being almost undistinguishable from an open structure.

Next, the structure under study has to be discretized in pieces, which we called slabs, that are invariant in the z-direction, see figure 3.1. This might, at first, seem a limitation of the method but by choosing the discretization step arbitrarily small every continuous structure can be approximated as closely as desired. It is immediately clear that eigenmode expansion is perfectly suited for structures as the one in figure 2.9. Within each slab the first N (complex) eigenvalues of the vectorial Helmholtz equation corresponding with the refractive index distribution within the slab have to be determined. These eigenvalues are represented by the symbol β_i and are called *propagation constants* in a waveguide context. Using the boundary conditions for electromagnetic fields at a refractive index interface, for each eigenvalue an accompanying eigenmode, or mode profile, can be constructed. These eigenmodes, shortly modes in what follows, are the natural ways, in which light can propagate in a waveguide. When all the energy is within one mode in a certain point in a (perfect) slab, it stays like that within that slab.

When working with reciprocal systems, see appendix B, each positive eigenvalue will have a negative counterpart, which will be the eigenvalue corresponding with a backward-propagating mode. Note also that when constructing an eigenmode from an eigenvalue that the mode profile is undetermined by a factor that can be freely chosen. This factor, however, is mostly defined by a normalization for mathematical purposes.

When the first N modes of each slab are completely determined, the slabs have to be put together somehow. At each interface between slab I and slab II, 4 scattering matrices are to be determined, namely $T_{I,II}$, $T_{II,I}$, $R_{I,II}$ and $R_{II,I}$. Element (i, j) of $T_{I,II}$ is the transmission between (forward-propagating) mode j in slab I and (forward-propagating) mode i in slab II, after passing the interface $I - II$, and is a complex num-

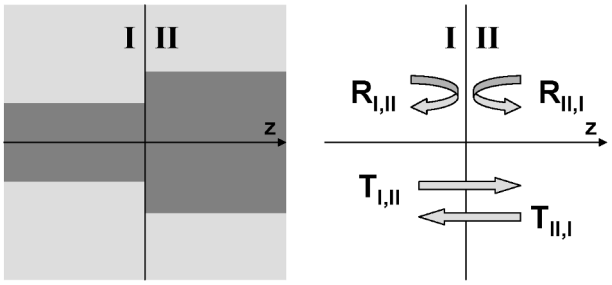


Figure 3.2: Scheme to clarify the 4 scattering matrices.

ber. And element (i, j) of $R_{II,I}$ is the transmission between (backward-propagating) mode j in slab II and (forward-propagating) mode i also in slab II, after hitting the interface $II - I$, see figure 3.2. The propagation within a slab (between two interfaces) is described by a propagation matrix, which is a diagonal matrix with element (i, i) being $exp(-j\beta_i L)$ where L is the distance between the two interfaces. Each slab has two propagation matrices, one for the forward-propagating and one for the backward-propagating modes.

An entire structure can be described by a total scattering matrix, which can be constructed from the various scattering (both R and T) and propagation matrices. This scattering matrix will connect forward- and backward-propagating modes of both the input and the output waveguides. Two different schemes can be employed to derive this scattering matrix, see [82]. One of the elements of this scattering matrix will describe the transmission between the forward-propagating ground mode of the input and the output waveguide. Immediately the advantage of the eigenmode expansion method (EEM) when calculating tapers becomes clear. The value of this element, or rather the modulus of its square, directly results in the efficiency of the taper under study. To improve the accuracy of an EEM calculation the number of modes can be raised until a convergence is reached.

Other simulation methods can be applied for studying optical tapers, like e.g. Finite Difference Time Domain (FDTD), the Beam Propagation Method (BPM) or Finite Element Methods (FEM). In all cases the field in the output waveguide has to be recorded and the overlap with the ground mode (or any other mode under study) must afterwards be computed. Compared with these methods, EEM calculates the taper efficiency 'for free'.

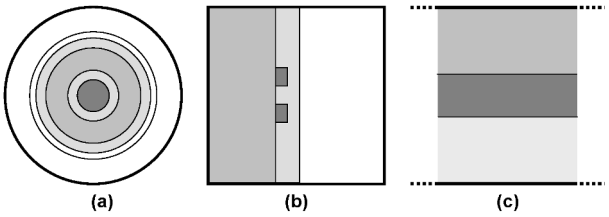


Figure 3.3: Three configurations implemented for CAMFR, the slab cross-section perpendicular to the optical axis is shown: (a) circular 3-D (b) cartesian 3-D and (c) cartesian 2-D. Different grayscales represent different refractive index, a dotted line means an infinite extension and electric or magnetic walls are in thick.

Finally, it has to be noted that EEM is a frequency domain method, meaning that the entire calculation is only valid for one specific value of the frequency (or wavelength). While this has advantages like the ability to work with dispersive media, it also has its limitations. A single EEM calculation does not provide time information, like the time it takes to charge an optical resonator, and no spectral information. While one FDTD calculation can give you a transmission spectrum, when working with EEM this has to be computed separately for every wavelength within the spectrum.

3.1.1 Camfr

For most of the work, an in-house developed eigenmode expansion method: CAMFR [83][84][85][86][87] has been used. This software is freely available (http://camfr.sf.net) under an open-source license and can be extended by anyone who would wish to do so. Originally developed for simulating VCSEL's, this method now contains three different mode-solving configurations: a circular 3-D, a cartesian 2-D and a cartesian 3-D mode solver, see figure 3.3. In the circular 3-D configuration a slab cross-section is a series of concentric circles, where each ring can have a different complex refractive index. In the cartesian 3-D a cross-section is composed of rectangles that can all have a different refractive index while in the cartesian 2-D you can only define a layer of refractive indices that extend to infinity in the other direction.

Most of the time we have used the cartesian 2-D configuration of CAMFR.

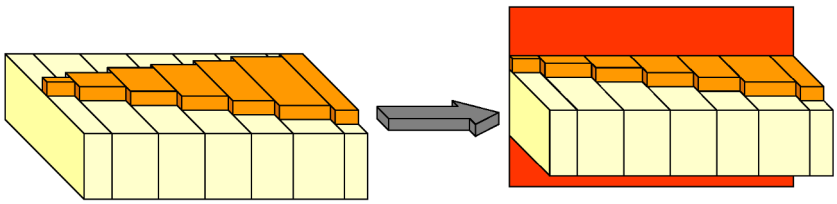


Figure 3.4: A mirror symmetrical structure can for calculation purposes be replaced by half of it.

3.1.2 FIMMPROP™

For a limited number of calculations, the commercial eigenmode expansion method FIMMPROP, distributed by PhotonDesign has been applied, sometimes to check important calculations, sometimes for the 3-D cartesian configuration, which is faster than the CAMFR implementation but less accurate because not all walls can be covered with absorbing boundary conditions. An optimization module KALLISTOS that closely interacts with FIMMPROP has been used for the (global) optimization of some 3-D structures that could be described using only a few parameters. With the number of parameters increasing, the optimization becomes less and less efficient and less useable.

3.1.3 Symmetry

If a structure has a mirror symmetry around a plane parallel to the propagation direction, then the simulation time can be significantly reduced by placing an electric or magnetic wall at the place of this mirror plane, see figure 3.4. This is due to the existence of even and odd modes with respect to this plane. If the optical axes of two consecutive slabs coincide, then there will be no mixing of these two mode types in the overlap integral at the interface. The use of a certain wall type will allow the existence of only one mode type, so depending on which mode type is required in a calculation, the corresponding wall type can be used in combination with half of the structure. In this case, only $N/2$ modes are needed to achieve the same accuracy as with N modes in the full structure calculation. Constructing the scattering matrix takes a time that is proportional to N^3, meaning that an eight-fold time gain can be reached using symmetry.

Note also that for some very symmetric structures, like a high-index rectangular waveguide embedded in a lower-index medium, two per-

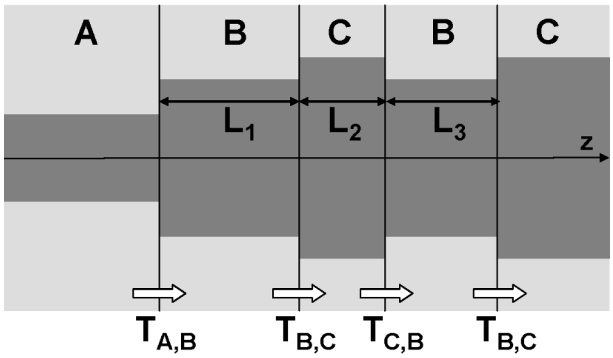


Figure 3.5: Unidirectional eigenmode expansion scheme.

pendicular symmetry walls can be used, which reduces the calculation to only a quarter of the structure.

3.1.4 Unidirectional Eigenmode Expansion

Instead of a bidirectional eigenmode expansion method , we have sometimes made use of the simpler, but approximate, unidirectional scheme.

In each slab of the structure, the first N eigenvalues β_i have again to be calculated and at each interface, see figure 3.2, 4 scattering matrices ($N \times N$) are determined. Contrary to the bidirectional case, only the scattering matrix $T_{I,II}$ is now further used.

For every waveguide section or slab, e.g. C, with a length L_2 a diagonal $N \times N$ propagation matrix P^{C,L_2} is derived by placing $exp(-j \cdot {\beta_C}^i \cdot L_2)$ as i^{th} diagonal element. Suppose, as in figure 3.5, that a certain structure consists of a cross-section A, followed by B with a length L_1, C with a length L_2, again B with L_3 and ends in C, then the total transmission matrix can be written as:

$$T = T_{A,B} \cdot P^{B,L_1} \cdot T_{B,C} \cdot P^{C,L_2} \cdot T_{C,B} \cdot P^{B,L_3} \cdot T_{B,C} \qquad (3.1)$$

and the ground mode intensity transmission equal to $|T(0,0)|^2$. Increasing the number of modes N will make the transmission converge, but not to the correct physical value. The unidirectional scheme does not take reflections at the interface into account and is only approximately correct in structures that will exhibit low internal reflection.

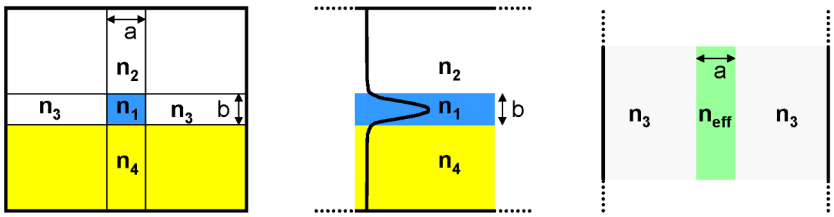


Figure 3.6: Illustration of the effective index method. **Left:** original slab cross-section. **Middle:** vertical middle stack to calculate slab ground mode. **Right:** final effective index distribution. A dotted line symbolizes an infinite extension.

3.2 Effective Index Method

Three-dimensional calculations take much more time than two-dimensional ones, making the former, nowadays, often useless for component design, especially in combination with an optimization, where many different structures have to be calculated. Therefore, two-dimensional calculations are often applied to demonstrate new principles and components.

To turn a real structure into an equivalent 2-D representation, one could think of taking an infinitely thin slice through the center of the waveguide core, parallel with the component plane. It turns out however that this is not the best choice and that the effective index method (EIM) is a much better option. This method was first described for rectangular optical waveguides by Hocker and Burns [88] in 1977. Instead of just horizontally cutting through the structure and taking the refractive index in that point, the structure is divided in vertical sections and in each section an optically averaged refractive index is defined in the following way.

The slab cross-section (perpendicular to the propagation direction) is divided in vertical sections, where the refractive index doesn't change horizontally, as in figure 3.6. The section that contains the waveguide core is transformed into a slab waveguide, by stretching it infinitely, and its ground mode is determined, resulting in a propagation constant β. Using the formula $\beta = 2\pi n_{eff}/\lambda_0$ an effective index n_{eff}, hence the name of the method, is attached to the propagation constant and can be thought of as the average refractive index that the mode encounters. In a next step this effective index is attributed to the entire vertical stack as if it was a real refractive index. During this procedure it is necessary to keep the field orientation in the original waveguide

in mind to set the correct boundary conditions every time. If the other vertical stacks can also be stretched to slab waveguides, the procedure is repeated in each stack, otherwise the refractive index (n_3) next to the waveguide core is taken, as in figure 3.6.

The usefulness of EIM for the simulation of 3-D components and in particular photonic crystals has been the subject of numerous studies, e.g. [89] and [90], in which the authors conclude that " *[...] a 2D calculation is not sufficient to describe a slab system for all frequencies. The two systems differ too strongly. However, using 2D calculations one might be able to identify configurations that could be suitable in slab calculations. Our results show that complete 3D computations have to be performed to obtain quantitative results for photonic crystal slab systems*".

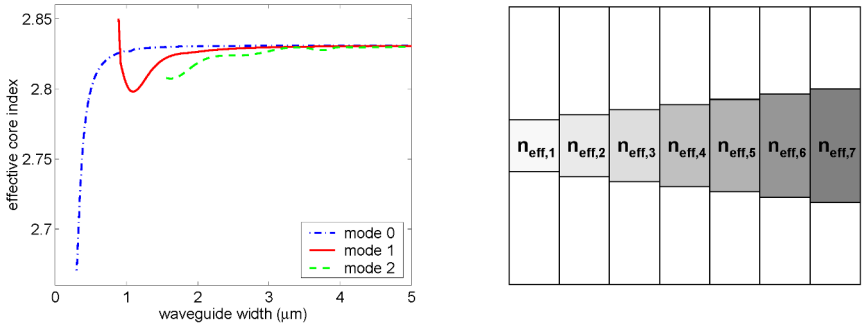
3.2.1 Enhanced Effective Index Method

Sometimes the EIM [91] is adapted slightly to more closely approximate a rigorous 3-D calculation. Instead of taking as effective index the value corresponding to the propagation constant of the 3-D slab waveguide, a value is chosen such that the propagation constant of the ground mode calculated in 2-D and 3-D is equal, for a certain waveguide width and wavelength.

For SOI wires with the standard layer structure (Si-substrate / 2 µm SiO_2 / 0.22 µm Si) and a silicon-only etch, this enhanced effective index for the ground mode is shown in figure 3.7(a) and is different for every possible waveguide width. For broader waveguides the enhanced EI is asymptotically identical to the conventional EI (value of 2.8305). A possible application in the simulation of a linear taper is shown in figure 3.7(b) with a different core effective index in every slab. Not only could such a configuration yield unexpected Fresnel reflections at the slab interfaces but as soon as more than one guided mode is involved, the enhanced EIM becomes unpractical, see figure 3.7(a). The enhanced EI for different modes can be totally different and can even lie at opposite sides of the conventional EI.

3.2.2 TE and TM

After applying the EIM a calculation becomes two-dimensional, meaning that the refractive index is invariant in the other, third direction. In this case the 6 coupled Helmholtz equations, one for each component of $\boldsymbol{E}$ and $\boldsymbol{H}$, decouple into 2 sets of 3 coupled equations, namely those for (E_y,H_x,H_z) and (H_y,E_x,E_z) with z the propagation, x the in-plane



(a) Enhanced effective index (b) Enhanced EIM

Figure 3.7: Left: for 3 modes the necessary core refractive index in a 2-D calculation to yield identical 2-D and 3-D effective indices as a function of the waveguide width. **Right:** a possible implementation.

and y the out-of-plane direction. When talking about 2-D calculations we will call the first case the transverse electric or TE polarization and the second the transverse magnetic or TM polarization. In a real 3-D structure this decoupling does not take place and modes will never be fully TE or TM but of a mixed polarization. However, mostly this mixed polarization strongly resembles TE or TM (the field components that should be zero are very small), so that one can speak of a quasi-TE or quasi-TM mode. In what follows, we will often speak of TE or TM in combination with 3-D structures, when we should actually use quasi-TE or quasi-TM. To avoid confusion between the conventional and the PhC definition of polarization or between the polarization of a waveguide mode and its polarization in an EIM calculation, the dominant electric field orientation will be indicated when necessary.

As in a 2-D calculation modes of opposite polarization will never couple at an interface, the modes of the non-intended polarization can be excluded for the calculation. This will again limit the number of modes in the calculation, like in section 3.1.3, and speed it up.

3.3 Optimization Algorithms

An optimization can always be seen as the minimization or maximization of a function. We will mostly talk about minimization, but the analogy for maximization should be straightforward.

For a function of only one argument, $f(x)$, long-established methods can be applied for minimization [92], some of which use the derivative $f'(x)$, if it can be computed. Also, if a function has a few arguments, $f(x_1, x_2, x_3)$, some of these methods can still be useful. However, as this number grows and the search space becomes multi-dimensional and has a lot of local optima, conventional minimization methods get into trouble and a new class of minimization algorithms must be applied [93]. Examples of these modern heuristics are simulated annealing (the first method to solve the traveling salesman problem in a practical way), the tabu search and genetic algorithms. As especially the last one is known for its ability to escape local minima during the course of the optimization, it was chosen in this work.

3.3.1 Genetic Algorithm

Genetic algorithms were first explored by John Holland [94] sometime in the mid-1960s when he tried to create artificial intelligence (AI) by mimicking the course of natural evolution in a computational way. More in particular, he let small computer programs evolve in ways that resemble natural selection to solve complex problems, like sorting a row of numbers. During the years, the method has found its way from the AI community to mathematics and engineering, where it is now used as an optimization tool.

In pseudo-code a genetic optimization algorithm looks as follows, where vocabulary taken from natural evolution is used:

Choose an initial population
Repeat
    Evaluate the fitness of each individual
    Select the best-ranking individuals to reproduce
    Let pairs mate at random
    Apply crossover operator
    Apply mutation operator
    Add new individuals to generation
    Remove the least fit individuals from the generation
Until a certain terminating condition

Individual: a certain set of parameter values $(x_1, x_2, \cdots, x_n)$ that completely describe a structure, in our case a taper structure. In the purest

genetic algorithm, not the parameters themselves but a binary representation is used, just as in biology not the proteins themselves are passed from a parent to its offspring but the DNA encoding of these proteins. We will stick however with the physical parameters (like a waveguide width) as this eases control during the course and the building of the algorithm.

Population: group of individuals. The initial population is most often randomly selected using very few limitations. This dumb creation of starting individuals is important because it allows a quick, although incomplete, scan of the total parameter space.

Fitness: the ability of an individual. In natural evolution the individual that is best fit for its environment will survive. In a genetic algorithm, the fitness function is the function under optimization. In our case, the ground mode transmission of the taper will be its fitness. If possible the fitness interval is chosen to be $[0 \ldots 1]$ because then the square or even higher powers of the fitness can be taken to better discriminate the best individuals.

Selection: the way individuals team up with each other before reproduction. In a genetic algorithm, this occurs in a random way but mostly the fittest individuals have a higher chance of being selected. The algorithm used in this work is known as Roulette Wheel selection (see later).

Reproduce: using two individuals to generate a few new individuals. Reproduction will be the result of both cross-over and mutation.

Mating: see Reproduce

Crossover: the operation whereby two individuals exchange information to create new individuals, in our case the mixing of two parameter sets

$$(x_1, x_2, \cdots, x_n) \;\&\; (y_1, y_2, \ldots, y_n) \tag{3.2}$$

to create two new sets

$$(x_1, x_2, \ldots, y_{n-1}, y_n) \;\&\; (y_1, y_2, \ldots, x_{,n-1}\,, x_n) \tag{3.3}$$

Various crossover operators exist, like uniform, one-point and two-point crossover. The two first have been used and will be explained

later.

Mutation: after crossover, the two new parameter sets are slightly altered, independently of each other. This is the equivalent of DNA changes due to copying errors or external factors like radiation. Mutation keeps the genetic algorithm from being stuck in local minima as soon as the genetic diversity drops within a population. The resulting individuals after the above crossover could look like this: $(x_1 + \Delta_1, x_2, \ldots, y_{n-1} + \Delta_2, y_n)$ and $(y_1, y_2 + \Delta_3, \ldots, x_{,n-1}, x_n)$ after mutation. Examples of mutation operators are flip-bit, boundary, uniform, gaussian, and other. Gaussian will be described in detail.

Generation: all individuals that are alive at a certain point during the algorithm. Contrary to natural evolution, parents and their offspring can be part of the same generation.

Removal: As we work with a fixed generation size, some individuals have to be removed from the generation at the end of each iteration. This dying is achieved by removing the individuals with the worst fitness values.

Terminating condition: As we want the optimization to end after a finite time, the algorithm is artificially stopped. This can happen after a fixed number of iterations, or when the fitness variation within a generation is smaller than a set value, or other conditions.

3.3.2 Implementation of the Genetic Algorithm

The choice of the genetic algorithm (which selection, crossover and mutation operator) and its implementation in the programming language Python were part of the subject of Peter Vandersteegen's master thesis [95].

Fitness

As an obvious choice for the fitness parameter, the taper efficiency, and more precisely the power ground mode transmission, was chosen. This is a real number between 0 and 1, which facilitates the rest of the algorithm. In what follows f_i represents the fitness of individual i.

Selection

At the start of each iteration during the course of the genetic algorithm, a few individuals are selected to reproduce. The higher the fitness, the more chance of being selected. In the current implementation, two non-identical individuals are chosen using Roulette Wheel selection, where an individual has a chance $P = f_i/f_{tot}$, with f_{tot} the sum of all fitness values within a generation, to be chosen. Mathematically, this is achieved by building a list with all the fitness values ordered in increasing order

$$(f_0, f_1, f_2, \ldots, f_n) \tag{3.4}$$

and converting this list into another one

$$(f_0, f_0 + f_1, f_0 + f_1 + f_2, \ldots, f_0 + f_1 + f_2 + \ldots + f_n = f_{tot}) \tag{3.5}$$

, which we call

$$(F_0, F_1, F_2, \ldots, F_n) \tag{3.6}$$

Then a real number F_{random} between 0 and F_n is computed using a random number generator and the index j for which $F_{j-1} < F_{random} \leq F_j$ is determined. The individual corresponding with the fitness f_j will finally be selected.

This procedure is equivalent to assigning each individual a certain section on a roulette wheel, with the size of the section proportional to its fitness. After spinning the wheel and determining in which section the needle comes to rest, the corresponding individual is chosen.

Crossover

When two individuals, the parents, have been selected for reproduction, their parameters have to be intermixed to create two new individuals, the children.

A first method for this intermixing is the *uniform crossover operator*, which is illustrated in figure 3.8. This operation can only be applied if both parents have the same number of parameters, which will always be the case in this work. For every parameter of parent 1, a 1 or a 2 is randomly generated with both 50% chance (symbolized by tossing a coin). If the outcome is 1 the parameter value becomes the first parameter of child 1 if not it goes to child 2, with the other child always getting the value of parent 2. This process is repeated for every parameter of the parents. The advantage of this procedure is the efficient scanning of the entire parameter space, as long as there is large genetic diversity

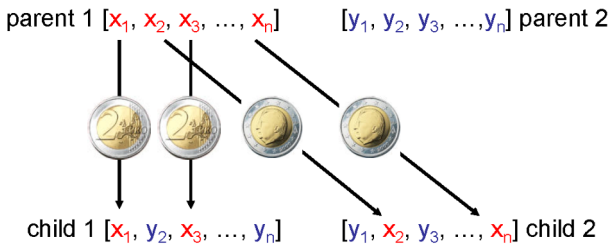


Figure 3.8: Illustration of the uniform crossover operator.

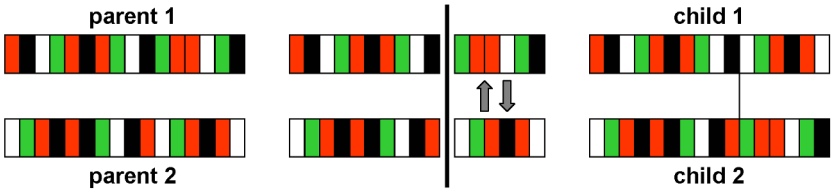


Figure 3.9: Illustration of the one-point crossover operator.

within a generation. The disadvantage is that children of widely differing parents, will lie far apart and far from the parents in the parameter space.

A second method to mix the genes of the 2 parents is the *one-point crossover operator*. We again assume that both parents have the same number of parameters, which we call N. An integer number n which lies between 0 and N $(0 \leq n \leq N)$ is randomly generated. The first n parameter values of parent 1 become the first n of child 1 and the same for parent and child 2, while the last $N - n$ parameter values of parent 1 become the last $N - n$ of child 2 and vice versa. This process is schematically shown in figure 3.9, where different grayscales represent different values.

The two-point crossover operator would generate two non-identical numbers between 0 and N and the resulting three gene pieces would alternatingly go to one of the two children. Intuitively it should be clear that an n-point crossover with n getting close to $N/2$ will be similar to a uniform crossover.

The advantage and disadvantage of a one-point crossover are opposite to a uniform crossover, namely the children resembling the parents more and a worse scan of the parameter space.

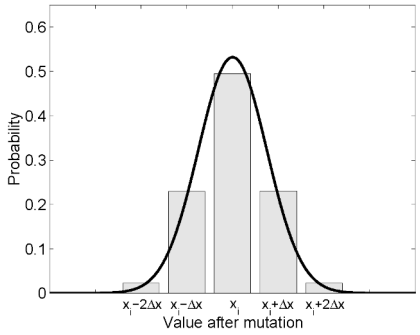


Figure 3.10: Illustration of the discretized and non-discretized gaussian mutation operator.

Mutation

After generating two children using the crossover of two parents, these children are slightly altered. This mutation brings genetic diversity within a population. At the beginning of a genetic algorithm when the genetic diversity is still large, this mutation enters new values into the algorithm (when working with float values) and slightly enlarges the search space, while at the end, when the diversity is very small, mutation is responsible for the identification of a local optimum.

We have made use of only the gaussian mutation operator. Every parameter value x_i of a set $(x_1, x_2, \ldots, x_n)$ is mutated independently. As input the gaussian mutation operator takes a number that determines the (averaged) size of the mutation and that we call Δx. As a first step a normally distributed random number x_i' is generated using the gaussian function completely determined by $\mu = x_i$ and $\sigma = 3/4 \cdot \Delta x$. In a simple non-discretized mutation x_i is now replaced by x_i' as the parameter value and the process is repeated for the other values as well. In a discretized mutation however an extra step is required. This kind of mutation is applied when a certain parameter can not take all possible values but only a discrete number like a number from the row $(0.3, 0.4, 0.5, 0.6, 0.7, \ldots)$. In the extra step x_i' is rounded off to the closest allowed value. If Δx is the step in the row of allowed values and if in the mutation operator $\sigma = 3/4 \cdot \Delta x$ then there is a probability of around 50% that the value remains unchanged, 46% that it changes with $\pm\Delta x$ and 4% that it changes with $\pm 2\Delta x$, see figure 3.10.

Often certain parameters under optimization can not take all possible values, not only because of discretization but also because of bound-

ary restrictions. Crossover will respect these boundary restrictions as it starts from allowed values and only switches these between individuals. Mutation, however, can push a value out of its allowed interval, which has to be avoided. We have used the simplest correction step, which gives a value that is mutated out of its interval the value of the crossed boundary.

More extensive descriptions of genetic algorithms and various other selection, crossover and mutation operators can be found in [93] and [50].

3.3.3 Sequential Partial Derivative Descent

In one instance, we have used an optimization algorithm that we have not found described in the literature[1] and that we have baptized the 'sequential partial derivative descent' algorithm, as it can be considered as a variation of a 'gradient search' or 'steepest descent' optimization.

Consider a function with a large number of arguments:

$$f(x_1, x_2, \ldots, x_n) = f(\boldsymbol{x}) \tag{3.7}$$

that we want to minimize. If looking for a global minimum, a genetic algorithm would be a good choice. If we know, however, that the minimum is close to a certain known set of starting values,

$$\boldsymbol{x}_0 = (x_{1,0}, x_{2,0}, \ldots, x_{n,0}) \tag{3.8}$$

(first subscript = dimension, second subscript = iteration step) then a genetic algorithm would be overkill by scanning the entire parameter space and a local minimization would be recommended. A well-known local optimization method, also useful when optimizing a multi-dimensional problem is steepest descent. It starts from the known value $\boldsymbol{x}_0$ and iterates towards the local optimum using the formula:

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - \gamma \cdot \boldsymbol{\nabla} f(\boldsymbol{x}_i) \tag{3.9}$$

with γ a small positive number and $f(\boldsymbol{x}_0) \geq f(\boldsymbol{x}_1) \geq f(\boldsymbol{x}_2) \geq \ldots$. If the function f is not analytically known the derivative must be calculated numerically, which requires $2 \cdot i$ function evaluations for every step in the iteration. An example of a steepest descent search is shown in figure

[1] PS: It is treated in [92] under the name 'successive line minimization'. Thanks Dominic.

3.11(a), where it is clearly visible that the search path is perpendicular to the contour lines.

Sequential partial derivative descent is also an iterative search process. Instead of calculating the gradient of the function, a partial derivative to every coordinate is calculated separately, by evaluating the function in a point $\pm s$ away from the starting point,

$$\frac{\partial f(\boldsymbol{x}_i)}{\partial x_1} = \frac{f(x_{1,i} + s, \ldots) - f(x_{1,i} - s, \ldots)}{2s} \tag{3.10}$$

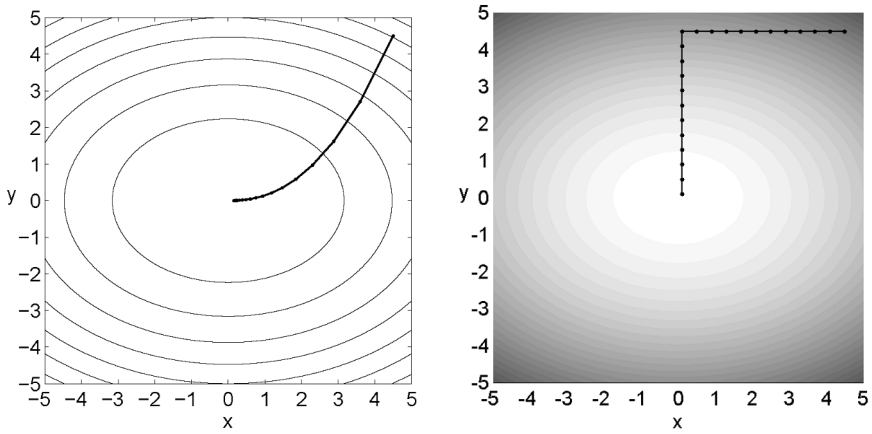
with s a small step value. Depending on the value of this partial derivative, the new best value becomes

$$\boldsymbol{x}_{i+1} = (x_{1,i} \pm s, \ldots) \tag{3.11}$$

or stays equal. This process is repeated for the same coordinate x_1 until no further improvement is reached, after which coordinate x_2 is improved, and so on.

When all coordinates have passed, the process restarts with coordinate x_1, then coordinate x_2, ... An advantage of this method is that as long as one can improve by changing the same coordinate, only 1 extra point has to be calculated as the other 2 points can stay in memory. An example of the sequential coordinate descent can be seen in figure 3.11(b), where the iteration across all coordinates has only happened once.

As a general optimization method, sequential partial derivative descent might be rather ineffective, especially the step-wise optimization of a single parameter. However, it has been implemented with a particular structure in mind and has proven to be a useful method to optimize that structure. More specifically it is used in the next chapter for the iterative optimization of a shortened parabolic taper.



(a) Steepest descent (b) Sequential coordinate descent

Figure 3.11: Illustration of steepest descent and sequential coordinate descent on the function $f(x, y) = x^2 + 2y^2$ with starting point $(4.5, 4.5)$. **Left:** function as contour lines, $\gamma = 0.1$. **Right:** function as grayscale, $s = 0.4$.

Chapter 4

Simulation Results

In chapter 2 a short overview of the scientific literature concerning tapers was given. Initially people have focused on the question which taper shape guarantees the least coupling from the ground mode away to other modes. This taper form is proven to be parabola-like and depends on the exact parameters of the coupling problem under study. In practice, knowing the perfect shape is not necessary as one uses parabolic or even linear tapers that open so slowly that the inter-mode coupling is negligible. Also theoretically, this perfect shape is not very useful, because a taper will never be infinitely long, and inter-mode coupling will always exist. When two modes in the same waveguide get excited, mode beating will result in a (length-)periodic transfer of power between the modes. Only for the lengths where all the power is in the ground mode, the perfect shape would indeed be perfect. For all other lengths, a different shape that generates a higher inter-mode coupling could, contradictorily, result in a better taper.

In this chapter we will follow the inverse approach. Instead of studying which taper shape generates the least coupling to other modes and then calculating what the minimum taper length is for adiabatic behavior, we will start from a given taper length and optimize the shape of the taper for this length. Of course, not all lengths will result in a decent taper but, by moving away from strictly adiabatic structures, novel taper shapes can be generated with this approach.

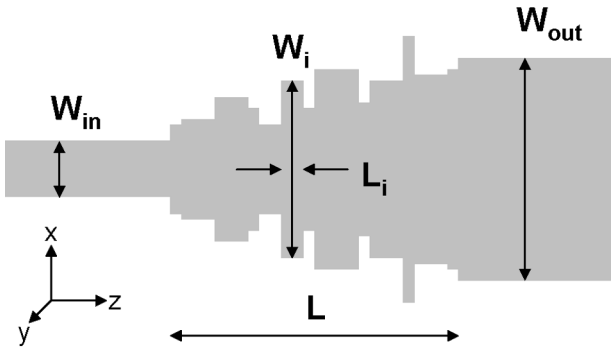


Figure 4.1: Scheme of the taper problem in its most general form. Only the waveguide core is shown.

4.1 Introduction

As we will mostly apply eigenmode expansion to calculate taper structures, which implies dividing these into z-invariant waveguide slabs, the construction of an efficient taper with a fixed length is equivalent to determining 'good' parameters for these waveguide slabs. Figure 4.1 shows a taper example in its most general form. Between the input waveguide with width W_{in} and the output waveguide with width W_{out} N waveguide sections each characterized by a width W_i and a length L_i are placed with

$$\sum_{i=1}^{N} L_i = L \tag{4.1}$$

as length requirement. For rigorous 3-D calculations the vertical light-confining layer structure has to be completely taken into account. In 2-D simulations, after an effective index transformation, the simulation will look like figure 4.1, where only two refractive indices are used, one for the core and one for the cladding.

The various sections of this chapter tackle the optimization problem, each using a different approach. The initial conditions can vary, sometimes starting from a parabola and trying to improve it, sometimes starting from completely random configurations. Also the parameters altered during the optimization can vary: at one time the 2 parameters (W_i, L_i) describing a waveguide slab are directly optimized, at another time the parameters describing a continuous envelop function are altered. For reasons of speed, rigorous 3-D calculations are not always feasible and approximations to the model that increase calcu-

lation speed are applied. The effective index transformation is applied most of the time, while some of the 3-D calculations have been obtained by relying on a unidirectional scheme. Finally, also the optimization algorithm applied in the various sections is not always the same.

The polarization considered throughout this chapter is the one where the dominant (transversal) electric field lies in the x-direction, see figure 4.1, and is mostly called TE in literature. This choice is a direct result of the substrate losses in a SOI wire. For the SOI layer structure we have used, the substrate losses for TE are smaller than for TM. The use of a thicker silica buffer layer could change this situation, which offers some benefits, but is technologically difficult.

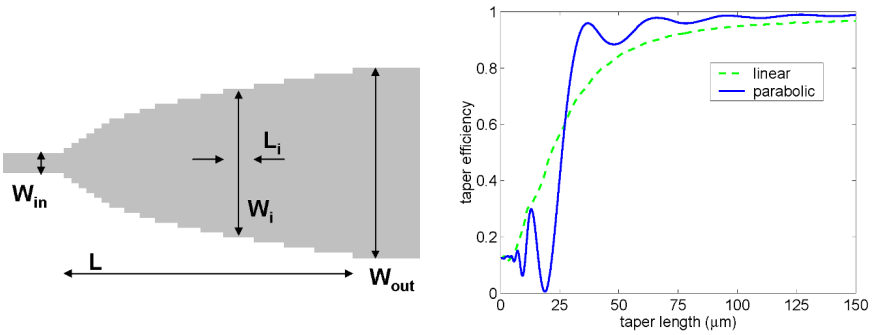
4.2 Modified Discretized Parabolic Taper

In this section the structural optimization problem presented in figure 4.1 is tackled by starting from a well-known and good-performing geometry: the parabolic taper, see also appendix A. Throughout this paragraph we will use an input waveguide width of 0.56 µm and an output waveguide width of 10.0 µm. For an eigenmode expansion calculation the continuous taper form has to be discretized into z-invariant waveguide slabs as in figure 4.2(a). If this number of slabs is chosen high enough the calculation gives the same result as for the original smooth structure. For this particular case 200 steps (and a symmetry wall) was enough to guarantee no residual effect of the discretization. As a first step the taper efficiency is calculated as a function of the length of the taper. This efficiency will be the butt-coupling efficiency (input and output joined together without a taper) for very short lengths and will be 1.0 or 100% for very long lengths. In between (see figure 4.2(b)) we see an almost monotonously increasing curve for the linear case and a periodic and more irregular curve for the parabolic taper. As theoretically derived, the parabolic taper reaches adiabaticity for shorter lengths than a linear taper. The observed periodicity (around 30 µm) in the efficiency for longer taper lengths corresponds very well with the (double) coupling length (31.2 µm) between the ground mode and the next even mode:

$$L_{coupling} = \frac{2\pi}{\bar{\beta}_0 - \bar{\beta}_2} \tag{4.2}$$

with

$$\bar{\beta}_i = \frac{\int_0^L \beta_i dL}{L} \tag{4.3}$$



(a) Discretized parabolic taper (b) Linear and parabolic taper efficiency

Figure 4.2: Left: parabolic taper discretized using a fixed width (x) step. **Right:** transmission efficiency of a linear and a parabolic taper, 200 waveguide slabs, W$_{in}$=0.56 µm, W$_{out}$=10.0 µm, λ=1.55 µm.

where β_i is the propagation constant of the i^{th} mode and L the total taper length. The first two large efficiency peaks (37 and 66 µm) will be used as starting points for the optimization, and discretized on a fixed x-step (transversal in-plane direction) with 60 sections.

As optimization algorithm the one described as sequential partial derivative descent in the previous chapter is used. One starts with the first taper section from the left (W_1, L_1) and calculates the transmission of the taper with $W_1' = W_1 \pm dx$. If one of these changes brings a gain to the transmission, W_1 is changed accordingly. Then W_1 is probed one dx-step further to see if further gain can be achieved until W_1 ends up in a local maximum. The same process is iterated for W_2, W_3, ... until one reaches the right side of the taper. Due to the changes to some of the other widths of the taper, W_1 is now no longer guaranteed to be placed in a local maximum, so the scanning with dx-steps restarts and can now lead to changes in the opposite width direction. Note, however, that with every change the taper transmission is improved. After W_1, all the other widths undergo the same procedure. This left-right scanning continues until a certain stop condition, which can be a maximum number of iterations, a minimum improvement due to the last iteration or a transmission threshold that is reached (e.g. 98%). In all the previous optimization steps the lengths L_i of the taper sections remained constant. After the left-right scanning iteration is stopped, all lengths are changed in the same step using the relation:

$$L_i' = a_s \cdot L_i \tag{4.4}$$

with a_s a shortening factor close to unity (e.g. 0.9). After the length change, the above width optimization change starts all over until the taper reaches the desired length.

The entire optimization procedure is characterized by the dx-step, the shortening factor a_s and the intermediate stop condition. These numbers can not be chosen arbitrarily and their value will determine the effectiveness and speed of the algorithm. For all optimizations dx is fix at 50 nm, which means 25 nm at every side of the section.

A first optimization run is visualized in figure 4.3 with as parameter values: dx= 50 nm, a_s=0.9, starting length = 66 µm and every left-right scanning is iterated six times. For lengths above 30 µm the consecutive left-right scans clearly improve the efficiency up to almost 100%, outperforming parabolic tapers. For shorter lengths, the improvements can still be as high as 20% but remain far away form the optimal efficiency. Interesting is the observation that after every length change, the efficiency of the contracted taper returns to the parabolic taper curve. This means that a parabolic taper with small random variations in the taper curve will in fact act as a parabolic taper. From an optimization point of view this returning is unwanted because it demonstrates that the various optimizations during one run are independent of each other and that width changes that are good for a particular length become random changes for the smaller taper length. The conclusion for this run is that the value of a_s is too small and should be closer to unity to guarantee that beneficial width changes stay good for smaller lengths.

Another a_s value, namely 0.95, is used in a second optimization run, which is shown in figure 4.4(a). Other settings are 31.32 µm as starting length and a maximum of 20 left-right scans. In contrast with the first run, length changes do no longer result in a return to the parabolic taper curve but remain much higher. This means that width changes beneficial for a certain taper length, still contribute to a higher efficiency when the lengths of the taper sections decrease by 5%. In the length region where the efficiency of a parabolic taper drops from around 80% to almost 0%, the modified parabolic taper has an efficiency above 95%, effectively halving the length of the shortest well-performing structure from 36 (parabola) to 18 µm (modified parabola).

Other optimization runs with different settings have been performed. For every taper length within these runs the highest achieved efficiency is shown in figure 4.4(b). For lengths larger than 15 µm high efficiency are obtained but below this length, values drop suddenly. We believe that this sudden drop is not intrinsic to the problem under

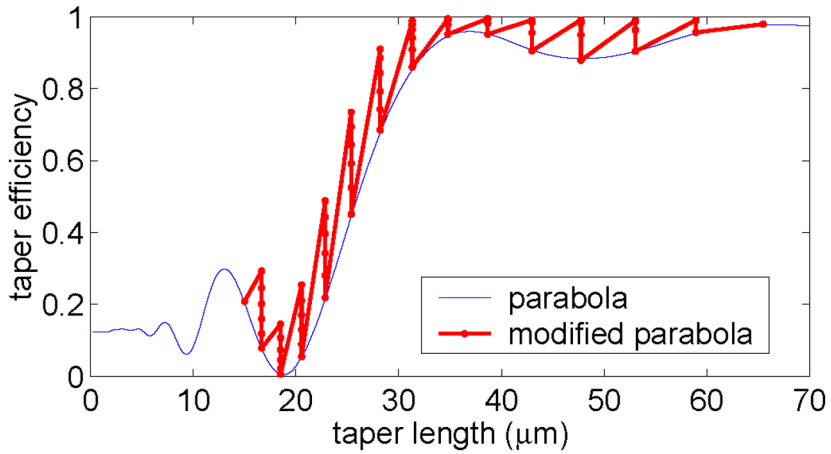
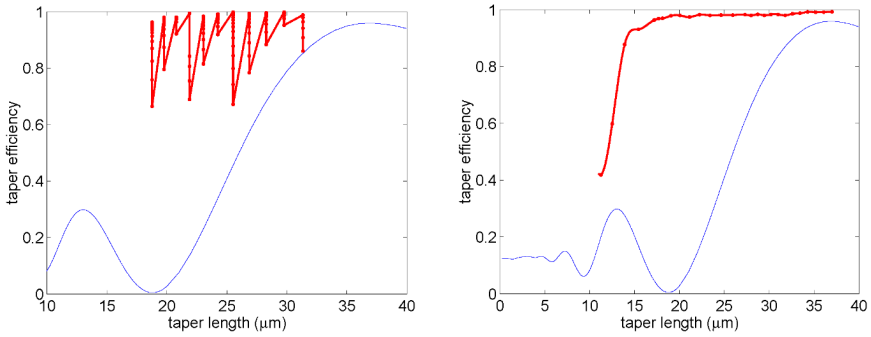


Figure 4.3: Efficiencies during first optimization run with unsuccessful parameter settings. Full line shows the efficiency of an unmodified parabolic taper.



(a) Successful optimization (b) Optimized values

Figure 4.4: Left: Optimization run with a_s=0.95, maximum of 20 left-right scan and a starting length of 31.32 µm . **Right:** Best efficiency value for every generated taper length. Full line shows the efficiency of an unmodified parabolic taper.

study but that for shorter lengths it becomes harder and harder to locate a good local optimum. With values for a_s even closer to unity and with more left-right scans the efficiency should also improve for shorter lengths. Because this strongly increases the necessary calculation time, this becomes impractical and has not been checked.

To prove that good width changes survive a length change and that both length and width changes are necessary to reach a good result, the minimum in efficiency for a parabola around 19 µm has been optimized directly. After an initial rapid improvement in efficiency, the result converged to a value of 57%, with no further improvement thereafter. This value is a lot better than a parabolic taper but stays far away from the one generated by the above optimization method, proving the necessity of the iterative length reduction.

Figure 4.5 shows the field plots of some optimized configurations, generated during the same run. Clearly visible is the birth of some features during the optimization proces, like the bump near the wide end of the taper and the collection of small features at the narrow end. The flat phase front near the end of the taper (excitation is from the left) is a visual hint for a good efficiency.

4.3 Ab Initio Block Tapers

Instead of starting from a known, well-performing structure as in the previous section, the optimization of the parameters describing the various taper sections, see figure 4.1, uses no prior knowledge in this paragraph. An iterative descent optimization is not useful any longer as no local optima are known in advance in the immense parameter space. To illustrate the size of this parameter space, consider a taper with 14 sections with a variable width, but a fixed length, and allow these lengths to change between 0.3 and 3.0 µm in steps of 10 nm. This allows for $270^{15} \approx 3 \cdot 10^{36}$ possible combinations. To look for good positions in such a huge multi-dimensional parameter space genetic optimizations are well-positioned, see section 3.3.1.

4.3.1 Two-dimensional Calculations

From 0.5 to 2.0 µm

As a first example the coupling between waveguides with widths of 0.5 and 2.0 µm in silicon-on-insulator (effective index of 2.83) is studied

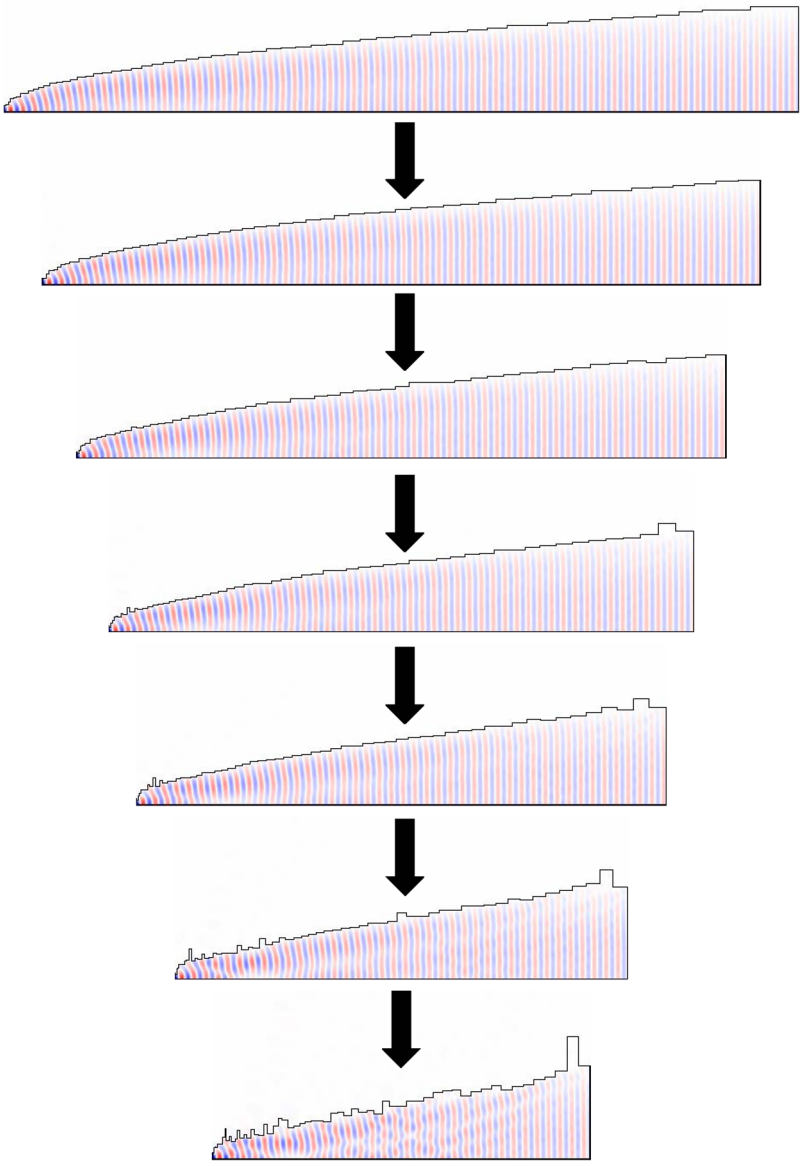
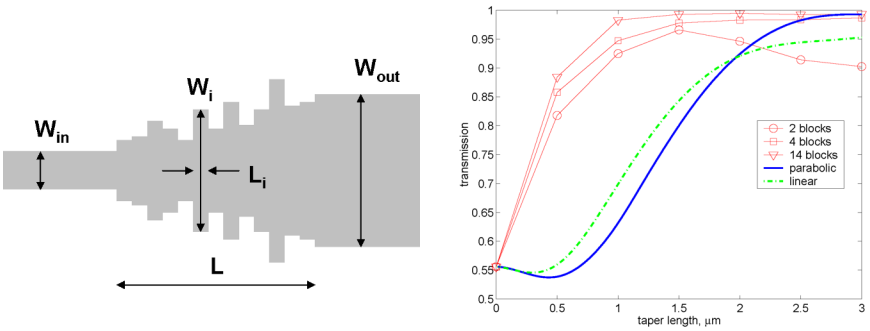


Figure 4.5: 7 field plots (H_y) of optimized tapers, only top half is shown, input and output waveguide not shown. Lengths from top downwards are 37.0, 33.4, 30.2, 27.3, 24.6, 21.1 and 17.7 µm.



(a) Block taper, equal lengths (b) Optimized efficiencies

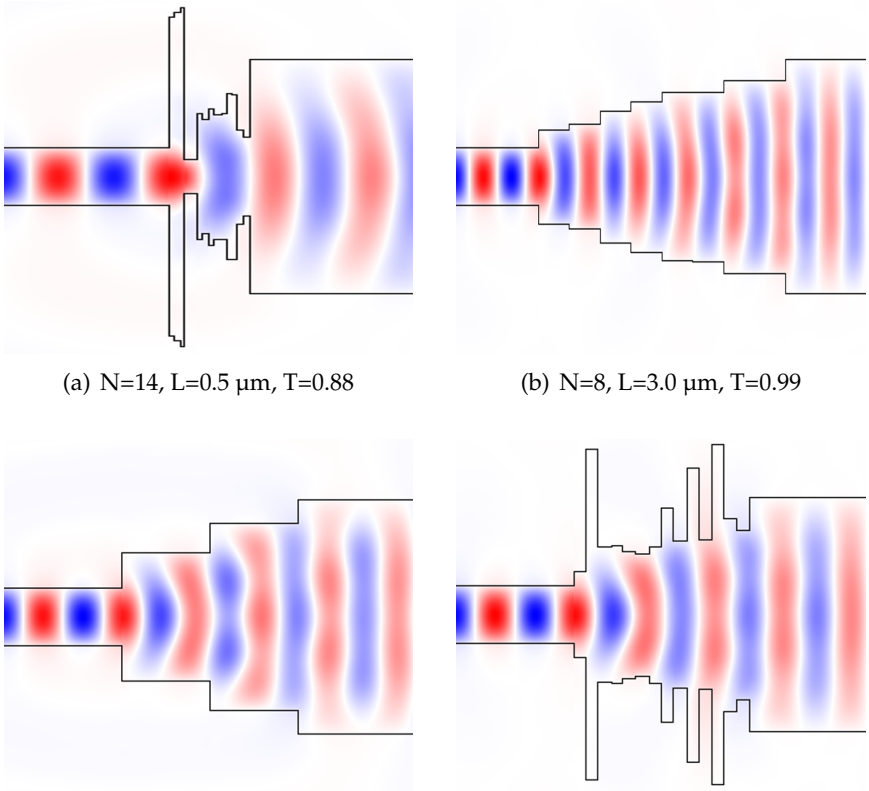
Figure 4.6: Left: Scheme of block taper as used in the optimization. **Right:** Efficiency as a function of taper length for various numbers of sections.

at a wavelength of 1.55 µm. The optimization is repeated for various lengths L of the coupling region and for various numbers of sections N involved, but the lengths of the sections is always considered fixed, namely $L_i = L/N$, see figure 4.6(a). Widths may vary between 0.3 and 3.0 µm in steps of Δx=10 nm.

As a start 250 individuals are randomly generated but only the best 100 end up in the first generation. The selection operator is a roulette wheel selection with the ground mode power transmission as fitness and selects two individuals. Cross-over is achieved by a uniform cross-over and the mutation operator is a discretized gaussian mutation on the two offspring individuals with σ=W_i and μ=$3/4 \cdot \Delta x$. The optimization is stopped as soon as the difference in fitness between the best and worst individual of the last generation reaches less than half a percent, which never needed more than 1200 generations.

For every taper length the number of taper sections is varied from 2 to 14 in steps of 2. Figure 4.6(b) shows some of the results, with every point being the end result of a separate genetic optimization. While a smooth linear and parabolic taper of length 1.5 µm respectively have an efficiency of 84% and 80%, the structure with 14 sections transmits 99% of the incoming power for an equal length. As is to be expected, the efficiency increases as the degrees of freedom of a structure increase, in this case meaning more sections.

Some of the resulting structures can be seen in figure 4.7, where the real part of the transversal (y-direction) magnetic field is shown when excited with the ground mode from the left. In figure 4.7(b) the en-



(a) N=14, L=0.5 µm, T=0.88

(b) N=8, L=3.0 µm, T=0.99

(c) N=2, L=1.5 µm, T=0.97

(d) N=14, L=1.5 µm, T=0.99

Figure 4.7: Field plot (real part of H_y) of some ab initio block tapers for coupling between 0.5 to 2.0 µm.

velope of the resulting structure strongly resembles a parabolic taper, which demonstrates the efficacy of the optimization method by reproducing the known best taper for this length. Surprisingly the structure in figure 4.7(c) with only 2 sections performs 13% better than the linear taper with the same length. The two tapers with 14 sections look highly irregular with strong width variations throughout the structure but also manage to significantly boost the efficiency. In the next chapter we will discuss some of these structures and provide some insight in the operation.

It can be expected that the tapers in figures 4.7(a) and 4.7(d) show a strong resonant behavior due to the various interfaces between low and high refractive index within these tapers and the accompanying reflec-

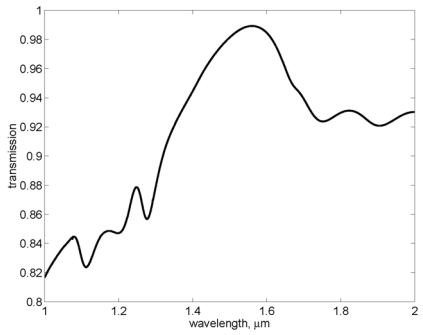


Figure 4.8: Transmission spectrum of the taper of figure 4.7(d)

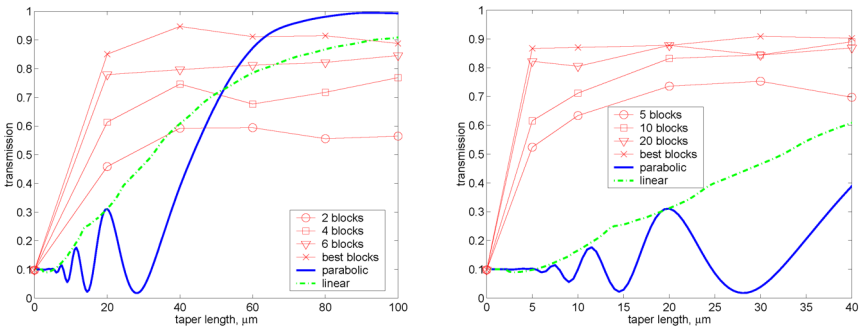
tions. However, when looking at the spectrum of the longest of these two structures, see figure 4.8, this does not show the strong fluctuations typical for a resonance but a relatively large wavelength region (1.5-1.6 µm) where the efficiency remains high (above 98%) and rather flat.

It is important to notice that a taper between two SOI waveguides with the above widths (0.5 and 2.0 µm) is already adiabatic when its length is at least 10 µm (both linear and parabolic). As the footprint of such a taper is modest the achievable gain in PIC area is almost negligible. This case should therefore be seen as an example that can be ported to other coupling problems where a much bigger gain can be achieved.

From 0.5 to 12.0 µm

A more practical coupling situation is the connection between a 0.5 µm and a 12 µm wide waveguide, which occurs when a vertical fiber coupler as in [66][67][53] is connected to a monomodal silicon-on-insulator waveguide. The mode width inside a standard single-mode silica fiber is around 12 µm and determines the width of the fiber coupler.

The same genetic optimization algorithm as in the above narrower problem is used with the exception of the possible widths, which now lie between 0.3 and 13.0 µm and can vary in steps of Δx=25 nm. As a test, next to the uniform cross-over operator, also a one-point cross-over operator is applied. For both lengths and widths the same cut position is used in the one-point cross-over meaning that the length and width of a certain parent section stay together during cross-over. To compare the two cross-over operators, the optimizations are run twice with all other parameters identical. Systematically, the results obtained with



(a) Coarse run (b) Finer run

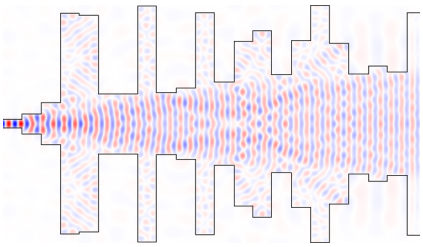
Figure 4.9: Optimized efficiencies for block tapers with various lengths and number of sections between waveguides with widths of 0.5 and 12 µm .

the uniform cross-over are slightly better than with the one-point cross-over and are shown in what follows.

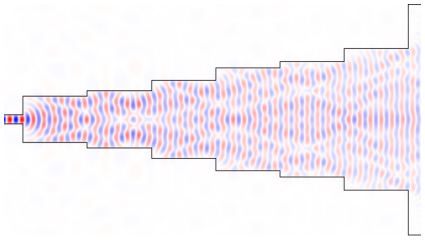
In a first run the taper length is swept between 0 and 100 µm in steps of 20 µm, and the number of sections between 2 and 14 in steps of 2. A selection of the results is shown in figure 4.9(a) together with the transmission of linear and parabolic tapers. As the complexity of a block taper increases the efficiency usually also goes up. This, however, is not always strictly the case as a genetic algorithm does not guarantee the global optimum. The line denominated 'best blocks' connects the best efficiencies, which appear depending on the taper length for 10, 12 or 14 sections. For lengths above 60 µm block tapers (up to 14 sections) function worse than conventional parabolic tapers.

A second run is more focused on the length region where a big efficiency gain can be achieved and uses more sections to construct a taper. Figure 4.9(b) shows these results, where 'best blocks' again shows the best efficiency regardless of the number of sections involved. From the figure it is also clear that, except for very short lengths, the efficiency is more or less independent of the length of the taper, but highly dependent on the number of sections of the taper. Block tapers with only a few sections can bring a high efficiency gain for short tapers, but very high efficiencies above 95% (-0.2dB) are never achieved for this coupling situation.

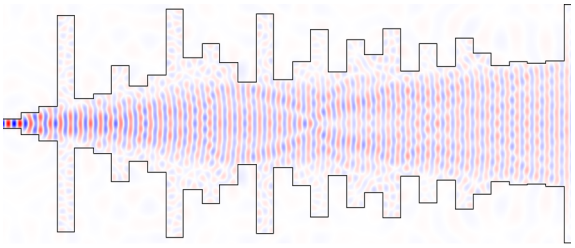
The field plots of some representative structures are shown in figure 4.10. With only few sections available, the taper forms a linear shape, like in figures 4.10(b) and 4.10(e), although the wide part of the taper



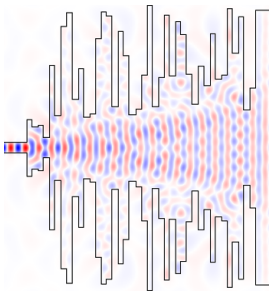
(a) N=20,L=20µm,T=0.88



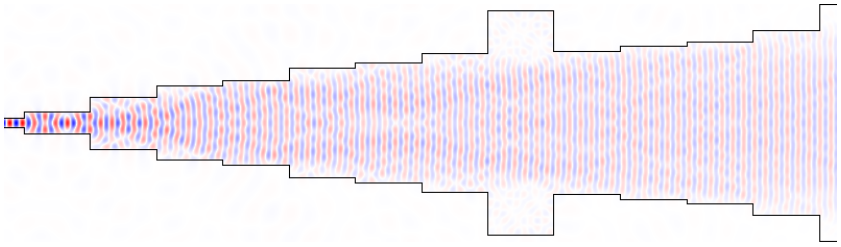
(b) N=6,L=20µm,T=0.78



(c) N=30,L=30µm,T=0.91

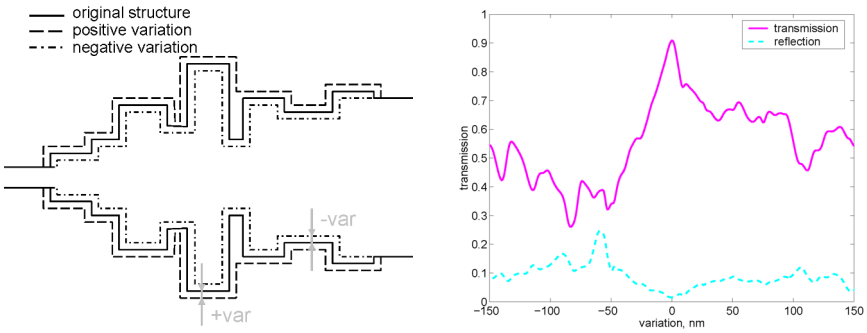


(d) N=40,L=10µm,T=0.87



(e) N=12,L=40µm,T=0.95

Figure 4.10: Field plot (real part of H_y) of some ab initio block tapers for coupling between 0.5 to 12.0 µm.



(a) Applied structure variation

(b) Effect on the transmission

Figure 4.11: The applied variation and its effect on the transmission of the structure in figure 4.10(c)

tends not to open up completely towards the wide waveguide. Tapers with more sections also show the linear shape but with some sections 'sticking out'.

The structure of figure 4.10(c) is used to study the sensitivity of the transmission to variations in the structural parameters. All structural parameters are changed simultaneously in the way represented in figure 4.11(a), where the variation procedure is shown on an example structure. While the widths of the input and output waveguides stay identical, the taper contour is redrawn somewhat smaller or bigger than the original contour. This variation can be seen as a simulation of over- or underexposure during the illumination step of the taper fabrication. We believe it is a good alternative for a sensitivity study where various parameters are changed randomly with, afterwards, a statistical analysis of the results. The transmission as a function of the variation on the structure of figure 4.10(c) is shown in figure 4.11(b). Around zero variation a narrow peak with a good efficiency is visible with a width of only approximately 10 nm. While this might seem a very low tolerance, it lies within the accuracy of the fabrication technology, explained in chapter 6. It is also to be noted that the structure under study is highly multimodal and the operation based on interference, which explains the small tolerance, and makes this taper a worst-case scenario example. The structures obtained for the coupling between 0.5 and 2.0 µm are more tolerant to structural variations.

Table 4.1: Details of 3D-optimized block tapers.

# blocks	lengths (µm)	widths(µm)	transmission
2	0.5 0.5	0.999 1.224	0.94
2	1.0 1.0	1.125 1.747	0.96
3	0.333 0.333 0.333	1.085 1.083 1.247	0.93
3	0.667 0.667 0.667	0.612 1.087 1.476	0.98

4.3.2 Bidirectional Three-dimensional Calculations

From 0.5 to 2.0 µm

Rigorous calculations in three dimensions take much more time than in two dimensions. For this reason the optimization of complex structures in three dimensions has not been studied. To couple two waveguides with widths of 0.5 and 2.0 µm only tapers containing 2 or 3 waveguide sections and having a length of 1.0 or 2.0 µm are optimized. As vertical layer structure SOI with 220 nm of silicon separated by a 2.0 µm layer of silica from the silicon substrate is applied. The etch depth is chosen equal to the silicon core thickness, a silicon-only etch. Since we are not interested in the very small influence of the substrate on the optical modes (in the forms of substrate losses) the silica layer is considered to continue indefinitely.

The 3-D eigenmode expansion solver FIMMPROP, see section 3.1.2 is applied to calculate each taper configuration, while the optimization tool KALLISTOS running on top of it, generates new taper configurations using a branch-and-bound optimization algorithm. No boundary absorption condition is selected but the width of the simulation area is chosen a few times larger than the total length, lowering the chance of reflections being reabsorbed. A central vertical symmetry wall is applied, which considerably speeds up a calculation. After the global branch-and-bound optimization, a local optimization that is also implemented in KALLISTOS improves the best global result. This combined approach takes several days for the simple tapers described here, turning more complex tapers impractical to study.

The details of the results of this 3-D optimization are given in table 4.1, with the order of the sections going from narrow to wide, and shown in figure 4.12. The data for the linear and parabolic tapers is taken from section 4.4.3, in which more details will be given.

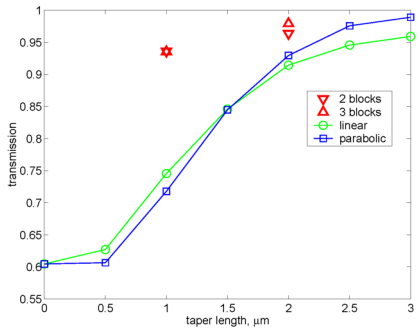


Figure 4.12: 3-D optimized efficiencies.

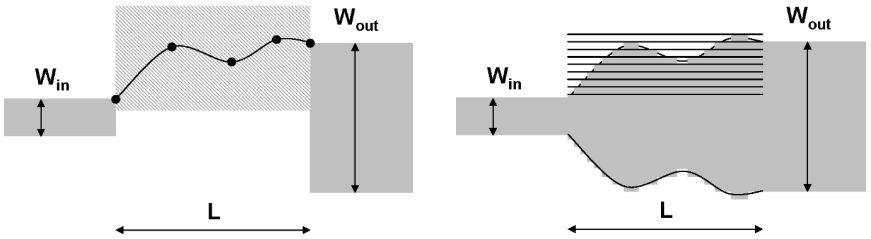
When we compare this figure with figure 4.6(b), the similarities are clear, which makes it plausible that with more sections, block tapers can also perform well when rigorously calculated. The butt-coupling efficiency in 2-D is only 55%, while in 3-D it is more around 60%. This can be understood by the lower confinement in 3-D, which pushes the mode somewhat more outside the waveguide core, and results in a better overlap with the ground mode of the 2 µm-waveguide, where the differences between 2-D and 3-D are much smaller.

The four optimized structures have been checked using CAMFR 3-D and the difference in power ground mode transmission between the results obtained using CAMFR-3D and those obtained using FIMM-PROP-3D was always smaller than 1%.

4.4 Spline Tapers

A third approach to optimize the various section parameters (W_i,L_i) of figure 4.1 is discussed in this paragraph. Instead of changing the parameters directly as in the above paragraph, an envelope function is now optimized from which the waveguide section parameters are derived afterwards.

Consider a taper with a fixed length L between two waveguides with given widths W_{in} and W_{out} as in figure 4.13 and focus on only one part bordered by the propagation/mirror axis. A virtual 'allowed box' is defined between the two waveguides with the lower limit defining the minimum waveguide width in the resulting taper (fabrication issue) and the upper limit defining the maximum waveguide width in the resulting taper, which affects the simulation area and the associ-

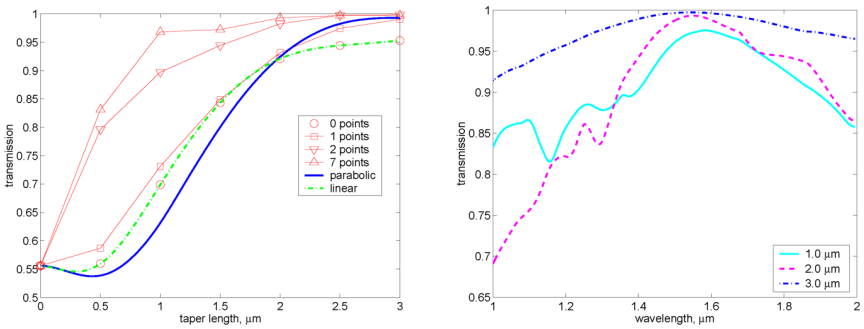


(a) Random points in allowed box (b) Connecting curve and discretization

Figure 4.13: Constructing a taper starting from randomly generated points within an allowed box, the connecting splines and the discretization on a fixed x-grid.

ated calculation time (simulation issue). Within this box a number N of points (z_i, x_i) is randomly generated that will define the shape of the taper. The points are connected (in order of ascending z-coordinate) using a cubic spline, which is a piece-wise third-order polynomial that satisfies the condition for a continuous zeroth and first derivative in each point. For the inner points these conditions completely define each piece of the spline. At the outer points, however, an extra condition is set by forcing the second derivative to zero, which has the advantage that a spline based on zero inner points coincides with a linear taper. A cubic spline with this boundary condition is called a natural cubic spline. Other possible boundary conditions in these points could be a certain value for the first derivative, in which case it is called a clamped cubic spline. The choice for a particular boundary condition, however, does not exclude certain structures as an extra spline point very close to a boundary point turns the condition in this point unimportant. When talking about the number of points a spline is based upon, the two outer points are not taken into account, as a result of which a zero point taper coincides with a linear one.

From this continuous curve, z-invariant waveguide sections need to be defined for use in eigenmode expansion. Various discretization schemes can be thought of, like a fixed section length or a fixed step in the width direction or others. While the former is easiest to implement mathematically, the latter is chosen here as it offers many advantages in combination with eigenmode expansion, because it greatly limits the allowed waveguide widths and therefore eases the recycling of already calculated quantities. From the predefined width step Δx and starting at W_{in} a grid in the x-direction is constructed, which sets the al-



(a) Optimized efficiencies (b) Some spectra

Figure 4.14: Left: Optimized spline taper efficiencies for various lengths, the number of points indicated in the legend. **Right:** Transmission spectra for spline tapers with 7 points, taper length indicated in the legend.

lowed section widths, see figure 4.13(b). By determining where the spline passes halfway between two grid lines, the lengths of the sections is defined. Together with some extra minor rules for local optima in the spline, this procedure unambiguously converts a spline in a well-matching discretized structure, ready for calculation. Other discretization algorithms have been compared with the above and as long as the step size is small enough, efficiency results converge to the same value.

4.4.1 Two-dimensional Calculations

From 0.5 to 2.0 µm

The genetic optimization follows the same scheme as in section 4.3.1 except for the cross-over, where now a one-point cross-over is applied to generate two new individuals. During the cross-over the two coordinates of a point stay together. Each point gets mutated using the same Gaussian distribution as before but with σ=7.5 nm and not discretized. Afterwards all points of a new individual are sorted with ascending z-coordinates to guarantee that the spline curve is a (single-valued) function. Even with all the points within the allowed box, the spline could still extend outside of it. If this is the case the offspring is rejected and the cross-over and mutation is iterated until a valid offspring is available. The discretization of the spline curve is based on a Δx= 12.5 nm, meaning that two consecutive waveguide sections differ by 25 nm. In

the next section it is shown that this step is small enough to consider a discretized structure as optically smooth.

Input and output waveguides have widths of 0.5 and 2.0 µm and calculated taper lengths go from 0 to 3 µm. For every taper length the number of points is varied between 0 and 7. Every marked point in figure 4.14(a) is the end result of a genetic optimization, except for the 0 points curve where there is nothing to optimize. The spectra of three of the structures based on 7 points are shown in figure 4.14(b), where it is clear that the bandwidth becomes smaller as the taper length decreases, leading to a sort of 'bandwidth-length quotient'.

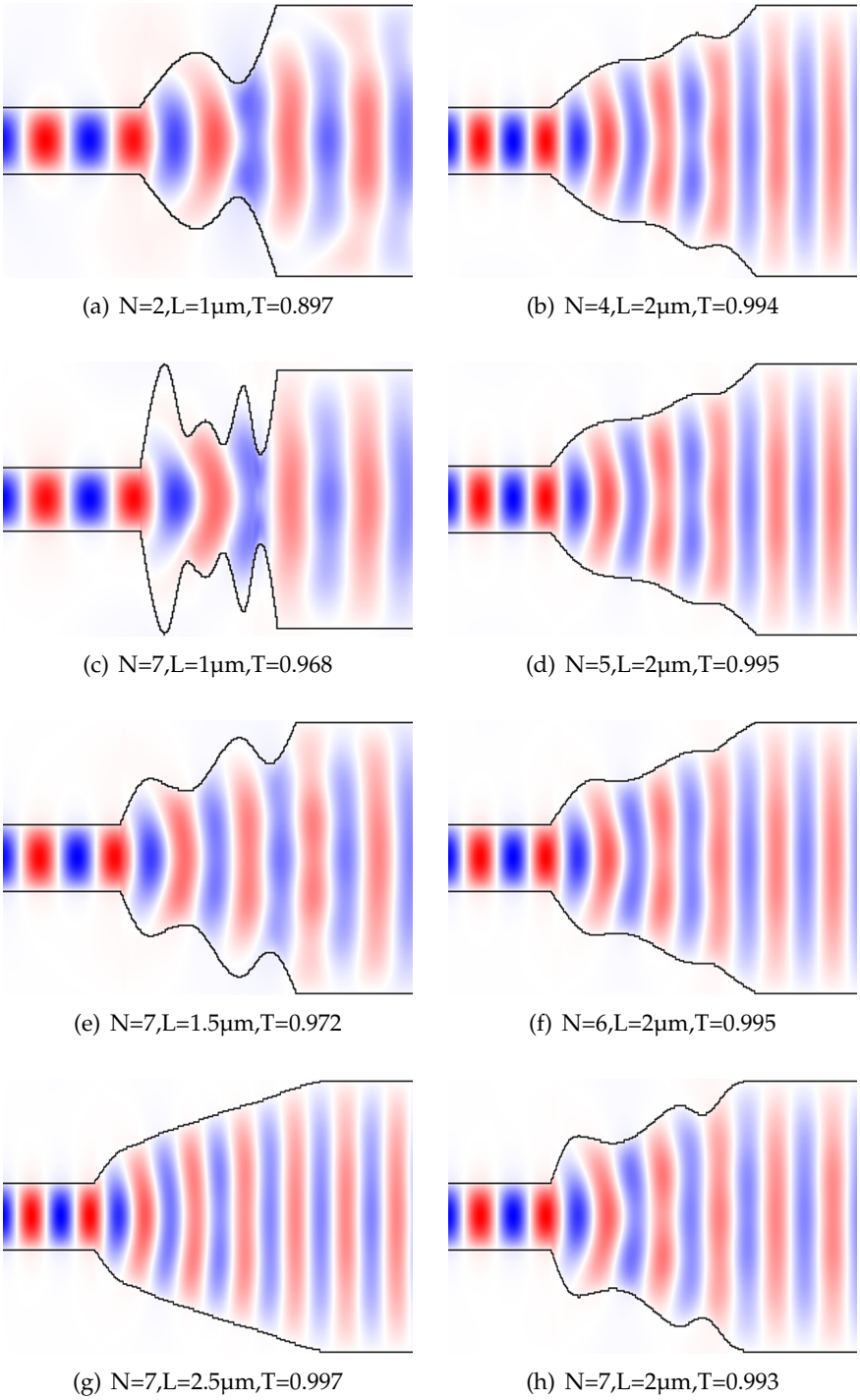
Again, it can be noted that a higher degree of freedom leads to an increasing efficiency. A genetic algorithm does not guarantee to find the absolute extremum of the function under evaluation, hence the shift away from the optimization wavelength 1.55 µm in one of the transmission spectra. Nevertheless, we have two reasons to assume our optimized structures approximate this (unknown) best structure very closely.

Firstly, the shape of structures with relatively long lengths are approximately parabolic, see figure 4.15(g) as an example. This is to be expected of a structure whose length approaches the condition for adiabatic operation.

And secondly, there is convergence within the results. For a fixed taper length but an increasing number of points, structures do not drastically change shape, as can be seen in the field plots at the right side in figure 4.15. As different structures are the results of totally independent optimizations this leads us to the conclusion that our structures converge to an optimal form that could be described by the spline approach using a large number of points.

From 0.5 to 12.0 µm

Using exactly the same genetic algorithm and discretization step as in the previous section, the coupling between 0.5 µm and 12 µm is studied. The discretization step Δx is doubled to 25 nm (50 nm width difference between two possible widths). To study the influence of this step size on the efficiency, it is swept between 6 and 500 nm for 4 randomly generated spline tapers. In figure 4.16(a) the convergence for small steps is clear, with 25 nm as upper limit before entering an unstable region. Although taking this upper limit introduces an error in calculated efficiency of up to 1%, this disadvantage must be weighed against the



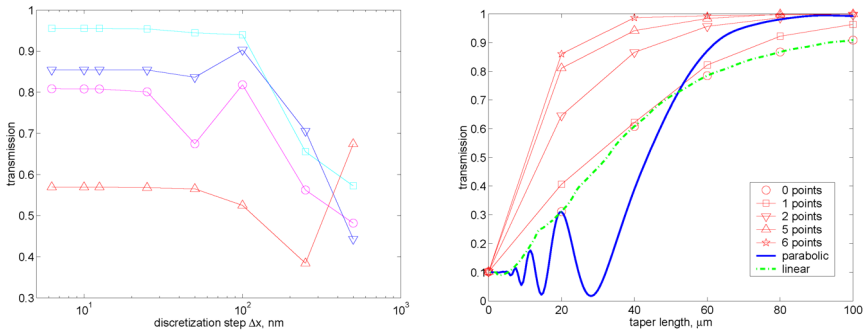
(a) N=2,L=1µm,T=0.897 (b) N=4,L=2µm,T=0.994

(c) N=7,L=1µm,T=0.968 (d) N=5,L=2µm,T=0.995

(e) N=7,L=1.5µm,T=0.972 (f) N=6,L=2µm,T=0.995

(g) N=7,L=2.5µm,T=0.997 (h) N=7,L=2µm,T=0.993

Figure 4.15: Field plot (real part of H_y) of some spline structures when excited from the left for the coupling between 0.5 and 2.0 µm. Figures not to scale.



(a) Efficiency versus Δx (b) Optimized efficiencies

Figure 4.16: Left: Convergence in efficiency for small discretization steps. **Right:** Optimized spline taper efficiencies for various lengths, the number of points indicated in the legend. Coupling from 0.5 to 12.0 µm.

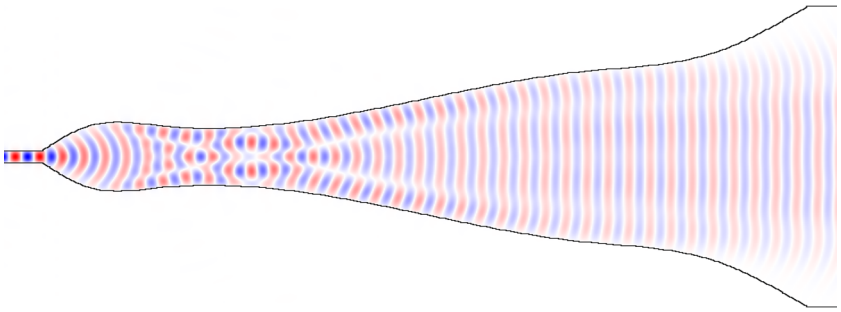
faster calculation times and less stringent computer memory (RAM) requirements.

Tapers lengths vary from 0 to 100 µm in 20 µm steps and 0 to 7 points have been used to construct the splines. The results are shown in figure 4.16(b) and some accompanying field plots in figure 4.17. When comparing figure 4.17(b) with figure 4.15(f), the similarity in taper shape is interesting. Instead of a gradually widening form, the opening of the taper is more a stepwise process, with 'plateaus' where the taper width stays constant. These plateaus also seem to be the origin of the lines of zero field, as in multi-mode interferometers. A similar behavior can be observed in figure 4.10(b).

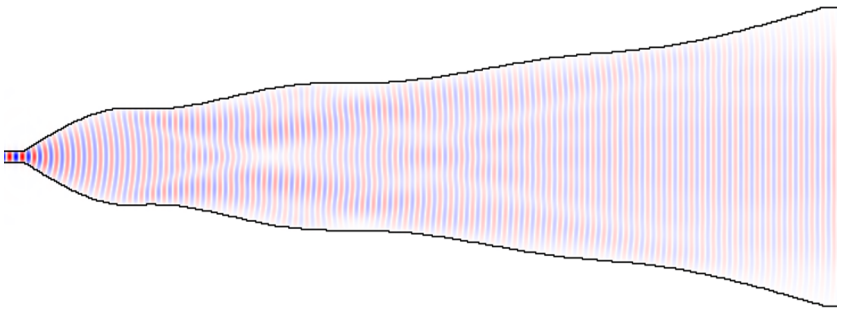
4.4.2 Unidirectional Three-dimensional Calculations

Because bidirectional three-dimensional eigenmode expansion calculations are still too time-consuming to be useful in combination with a structural optimization, a unidirectional approximation, as explained in section 3.1.4, is applied.

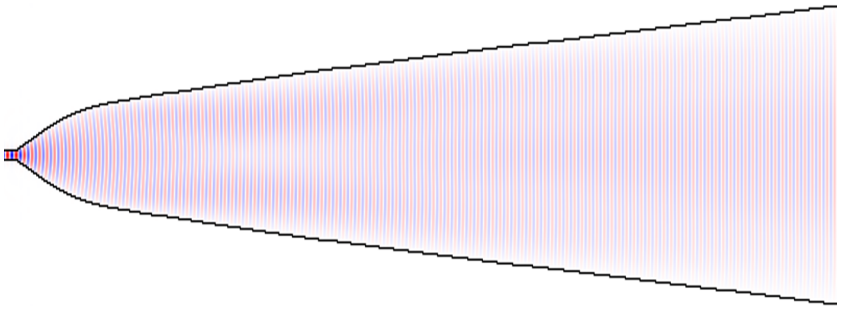
An advantage of starting from a continuous taper form and of the staircase approximation based on a fixed x-grid is that a certain waveguide cross-section can only be followed by the cross-section that is $2.\Delta x$ narrower or wider. This drastically reduces the number of possible transitions between cross-sections that occur during the course of an optimization.



(a) N=4,L=20µm,T=0.90

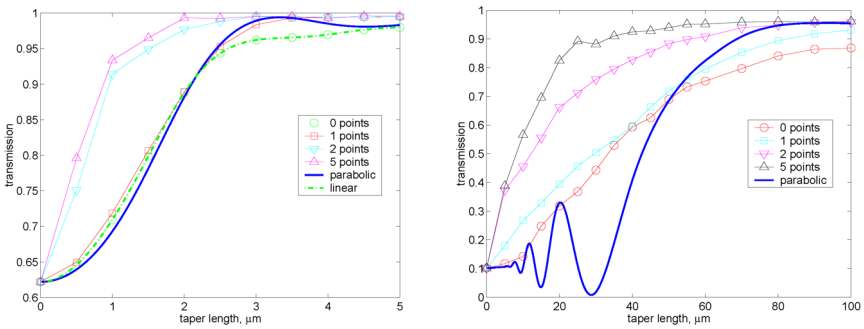


(b) N=6,L=40µm,T=0.99



(c) N=3,L=60µm,T=0.99

Figure 4.17: Field plot (real part of H_y) of some spline structures when excited from the left for the coupling between 0.5 and 12.0 µm. Figures not to scale.



(a) From 0.5 to 2.0 µm (b) From 0.5 to 12.0 µm

Figure 4.18: Optimized efficiencies for two coupling problems calculated with unidirectional three-dimensional eigenmode expansion.

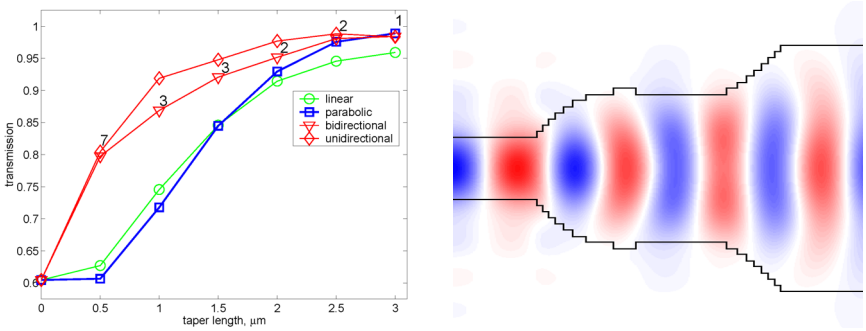
The propagation constants of every possible waveguide section and the complex overlap matrix between consecutive sections have been calculated once using FIMMPROP for a certain Δx and saved for repetitive use.

From 0.5 to 2.0 µm

The optimization algorithm is exactly identical to the one used to calculate the same coupling problem with the two-dimensional eigenmode expansion method. Figure 4.18(a) shows the resulting efficiencies as a function of taper length and for various numbers of spline points. Striking in this figure is the large difference in transmission between spline tapers based on one point and those based on two points, while adding even more points (up to 5) only seems to have a marginal effect as opposed to the two-dimensionally calculated spline tapers of figure 4.14(a), where there is also a big difference between 1 and 2 points but where adding extra points still brings a significant improvement. In the next chapter we will study this observation.

From 0.5 to 12.0 µm

All parameters of the genetic algorithm and of the discretization are identical to the two-dimensional coupling problem and results are shown in figure 4.18(b). These calculations are at the limit of what could be achieved using the available computer resources in a practical amount of time. Around half of the points in the graph (mostly for the tapers



(a) Efficiencies recalculated (b) A field plot

Figure 4.19: Left: Unidirectional efficiencies recalculated bidirectionally and the best results shown together with the number of points. **Right:** Field plot (CAMFR 3-D) of the taper with a length of 1.5 µm and based on 3 points.

with shorter lengths) are not fully optimized in the sense that there was not yet convergence within the final generation but the iteration was artificially ended after 25000 generations. This is probably the explanation for the curves not being as smooth as before.

4.4.3 Bidirectional Three-dimensional Calculations

As explained before, the optimized three-dimensional structures of section 4.4.2 do not take reflections into account. Because this approximation can be rather crude for short structures, it is necessary to check the results obtained with the unidirectional method afterwards with a correct model. The model applied here is the three-dimensional version of CAMFR [87], essentially a fully vectorial 3-D eigenmode expansion. A complete calculation of a structure consisting of around 20 different waveguide cross-sections takes a few hours of calculation time, making it impossible to use this model in a structural optimization as described above, and the reason why only the coupling from 0.5 to 2.0 µm has been recalculated.

From 0.5 to 2.0 µm

All structures of figure 4.18(a) with lengths between 0.5 and 3.0 µm and based on 1 to 7 points are recalculated, together with the accompanying (discretized) linear and parabolic tapers, see figure 4.19(a). For each length the most efficient structure is shown together with the efficiency

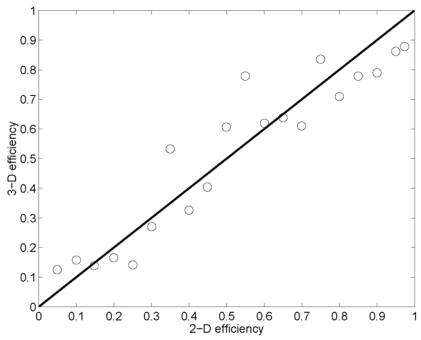


Figure 4.20: 3-D efficiency as a function of 2-D efficiency for a collection of block tapers.

calculated with the unidirectional method. One field plot is shown as an illustration in 4.19(b), with again a plateau in the taper shape. In all cases the correct efficiency is lower than the unidirectionally calculated efficiency, with the difference being higher for shorter lengths, which points to an increasing reflection in these structures. For taper lengths too short for adiabatic operation, the optimized structures always perform better, up to 20%, than parabolic and linear tapers of that length. While adding more degrees of freedom in all previous sections increased the efficiency, this seems not to be the case here. The structure leading to the highest efficiency is almost always based on 2 or 3 points. This means that the gain achieved by refining in the unidirectional optimization is below the error induced by the approximation of that method.

4.5 Necessity of Three-dimensional Optimization

From figure 4.19(a) it has become clear that an optimization with a unidirectional 3-D eigenmode expansion does not guarantee well-performing structures. However, if the effective index method, although an approximation, would be good enough to directly generate tapers that perform well in reality, than three-dimensional calculations would not be needed anymore in optimization, but only to verify the end results.

To check this assumption a genetic optimization in 2-D has been run for a block taper containing 8 sections with each a length of 0.5 µm to couple between waveguides 0.5 and 3.0 µm wide. During the optimization all parameters of each individual are stored and afterwards the in-

dividuals are determined that have efficiencies closest to 5%, 10%, 15%, ... up until 95%. For this list of tapers, the efficiency is recalculated using a rigorous 3-D eigenmode expansion method (FIMMPROP). In figure 4.20 the 3-D efficiency is shown as a function of the 2-D efficiency. If the effective index transformation would be perfectly accurate, all points would lie on the straight line. Although the global trend is certainly similar, there are deviations up to 20% between the two calculation methods. This leads us to the conclusion that a optimization in 2-D can be useful as a starting point but that an additional optimization in 3-D will still be necessary to obtain practical tapers.

4.6 Block versus Spline Tapers

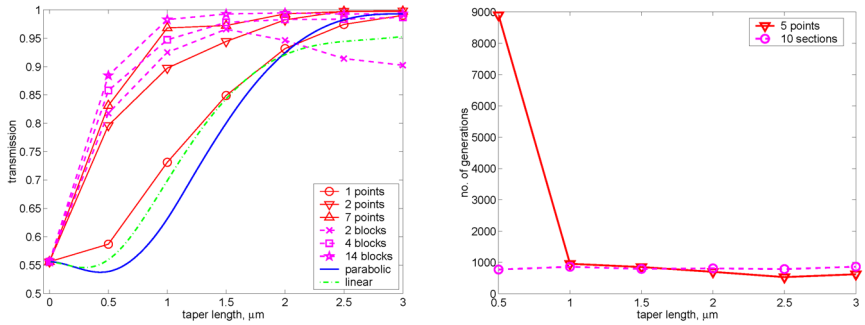
In this section we will constrain ourselves to the two-dimensional calculations and compare the block approach to the spline approach in terms of optimized efficiency and optimization time.

4.6.1 From 0.5 to 2.0 µm

In figure 4.21(a) the results for the coupling between 0.5 and 2.0 µm using both block and spline tapers is shown. In an attempt to compare these two methods for constructing a taper, a number of spline points is plotted together with the double number of sections. This way, the number of optimized parameters is equal because a point has two coordinates and only the width of a section is changed.

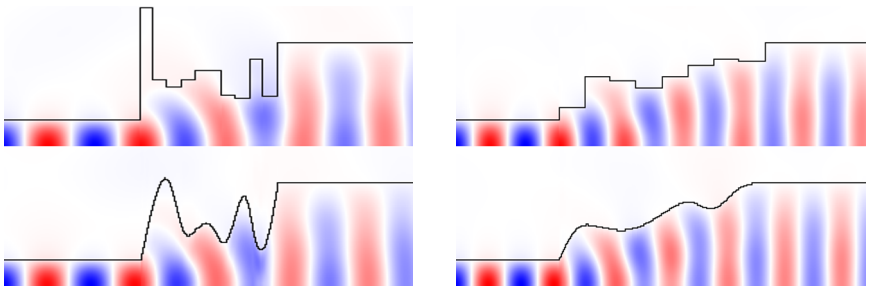
Up to a length of 2 µm, the block tapers perform better than their spline counterparts, above this length this is reversed. When looking at the number of generations required for convergence, see figure 4.21(b) this is around 1000 for both taper types except for the very short spline tapers, where the curve steeply rises. This increase is also found for spline tapers based on a different number of points.

Due to the differences between the two construction schemes, rectangular blocks versus polynomial functions, a certain structure cannot be exactly described by both. However, when comparing various final structures of the two cases, as in figure 4.22, some appear to be similar, not only in the form of the taper but also in the resulting effect on the field. Note that the length of a section is fixed and only the width is altered during the course of the optimization. Although this comparison is only qualitative and rather rude, it certainly points to a well-functioning of the optimization algorithm.



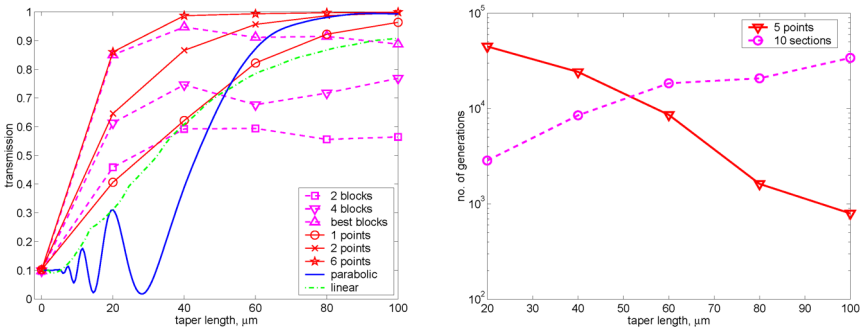
(a) Block and spline taper efficiencies (b) Number of generations needed

Figure 4.21: Comparison between block and spline tapers for the 2-D coupling from 0.5 to 2.0 µm.



(a) L=1µm, N_{spline}=7, N_{blocks}=10 (b) L=2µm, N_{spline}=7, N_{blocks}=8

Figure 4.22: Similar structures obtained using two different schemes to construct a coupler.



(a) Block and spline taper efficiencies (b) Number of generations needed

Figure 4.23: Comparison between block and spline tapers for the coupling from 0.5 to 12.0 µm.

4.6.2 From 0.5 to 12.0 µm

For the coupling between 0.5 and 12.0 µm a similar analysis has been performed. The optimized efficiencies in figure 4.23(a) show a different behavior with the spline tapers already having the best behavior for short lengths, starting at 40 µm. At 60 µm the 2-point spline taper even outperforms the 14-section block taper. Block tapers with a fixed number of sections need more iterations of the genetic algorithm when the taper length increases, while this is the opposite for spline tapers, see figure 4.23(b). Combining the two plots leads to the convenient conclusion that the fastest convergence also brings along the highest efficiency.

Probably the above comparison between a spline taper with N points and a block taper with $2N$ sections is not completely fair. For a very short taper, where we have no a priori knowledge about a good design, it might be, but for the somewhat longer tapers, there is a lot more information in the spline tapers than in the block tapers, think about a spline taper with 0 points. The longer the taper, the smoother it has to become for good coupling and spline tapers are smooth by design. The staircase step in spline tapers is a parameter defined independently from the number of points. For block tapers to reach an equal small width step between two adjacent sections, a huge number of sections would be required, probably too much for a practical optimization.

The block and the spline taper optimization in the case of coupling between 0.5 and 2.0 µm and the comparison between both, in two and

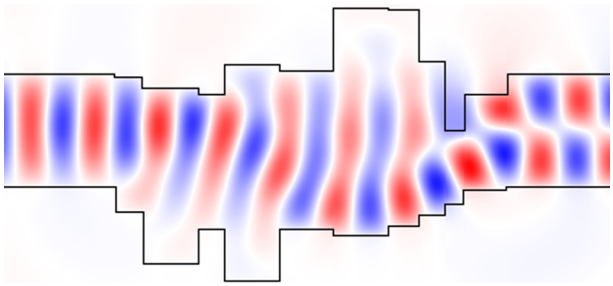


Figure 4.24: Structure optimized to couple from ground to first order mode and excited from the left.

three dimensions, has resulted in publication in the Journal of Lightwave Technology, see paragraph 1.6.

4.7 Other Objective Functions

Up until now, we have always optimized the same characteristic of a taper, namely the ground mode power transmission, $|T(0,0)|^2$. By choosing another objective function for the optimization, the resulting taper can be optimized for converting the incoming mode i into another outgoing mode j. Three different objective functions are described below, not in much detail but more to illustrate the strength of the applied optimization method.

4.7.1 Conversion to First Order Mode

As a first example of another objective function the transmission from the ground mode in the input waveguide to the first order mode in the output waveguide, or $|T_{0\rightarrow1}|^2$, is optimized. At the interface of two waveguides with the same symmetry axis, these modes do not overlap due to their different symmetries. As a consequence, the mirror symmetry around the propagation axis has to be broken to construct a coupler between these modes.

The genetic algorithm needs to be extended to not only handle the width and the length of a waveguide section, but also its shift with respect to a predefined zero line. Input and output waveguides both have a width of 1.0 µm and share the same symmetry axis, to which all shifts are defined. The coupler is built up out of 10 waveguide sections, with a length between 0.1 and 0.5 µm in steps of 25 nm, a width between 0.4

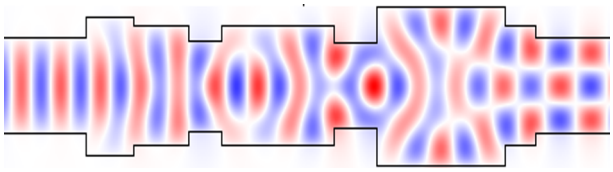


Figure 4.25: Structure optimized to couple from ground to second order mode and excited from the left.

and 2.0 µm in steps of 50 nm and a shift between -0.5 and +0.5 µm in 50 nm steps. The rest of the genetic algorithm is similar to the one used to optimize the block tapers described before.

The best structure, calculated in 2-D, has a length of 3.5 µm and converts 90.2% of the incoming ground mode to the outgoing first order mode, with 1.8% of the power still in the ground mode, and a total reflection of 2.2%. A field plot of this structure is shown in figure 4.24.

4.7.2 Conversion to Second Order Mode

Also the conversion from the ground mode to the next even mode, the second order mode is studied. As these two modes overlap at the interface of waveguide sections with the same symmetry axis, we can again limit the study to symmetric structures.

A structure comprising 7 sections, the length and width of which can vary between 0.4 and 2.0 µm in 50 nm steps, has been optimized, again in 2-D, for the same widths of the in- and outgoing waveguides.

The best structure has a length of 6.5 µm and converts 97% into the second mode with a residual 0.2% still in the ground mode, and a total reflection of 0.7%. A field plot of this structure is shown in figure 4.25.

4.7.3 Linear Combination of Mode Excitations

The conversion from a ground mode to a linear combination of various modes arises when one tries to couple from a photonic crystal waveguide to a broad dielectric waveguide. Let us reconsider figure 3.2 but with the left part not a dielectric but a (half-infinite) photonic crystal waveguide as in figure 4.26. Because one can cut various cross-sections within a photonic crystal waveguide, talking about the modes of such a waveguide does not make sense. However, as we have a periodic structure, Bloch modes can be defined. These are the natural propagation modes of light in such a structure and are determined by looking

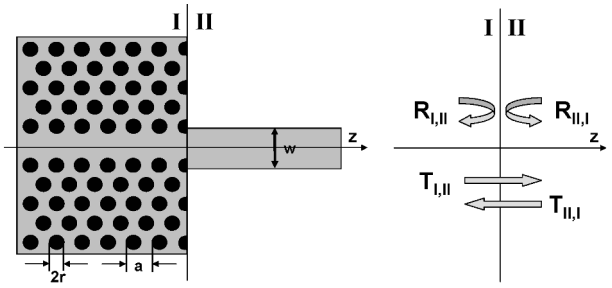


Figure 4.26: Definition of the various scattering matrices at the interface of a semi-infinite photonic crystal and semi-infinite dielectric waveguide.

for certain mode excitations that after passing one crystal period remain identical to themselves, except for a phase (and sometimes attenuation) factor. For every possible section within a crystal period, a Bloch mode can be decomposed into conventional modes of the waveguide structure of this section.

Imagine a semi-infinite photonic crystal waveguide interfacing with a semi-infinite dielectric waveguide as in figure 4.26. At the interface, scattering matrices, as defined in this figure, uniquely describe the transmission and reflection between the Bloch modes at the left and the waveguide modes at the right. The construction of the matrices $T_{II,I}$ and $R_{II,I}$ is described in [96] and $T_{I,II}$ and $R_{I,II}$ in [97]. Using these scattering matrices it becomes possible to calculate the transmission efficiency between the ground Bloch mode of the W1 photonic crystal waveguide into the ground mode of the dielectric waveguide.

The following parameter values are chosen to work out this problem: λ=1.55 µm, r = 0.3a, a = 0.470 µm and n = 2.7, not the effective index of the ground mode of the SOI slab, but somewhat lower because this gives a better match between the 2-D and 3-D band diagram, see also section 3.2.1. For these values the even Bloch mode is guided within the band gap and has a high group velocity, strongly resembling the ground mode of a dielectric waveguide. The mode and the band diagram is visualized in figure 4.27 with the normalized working frequency equal to a/λ=0.303 and k_z=0.259.

Because the Bloch mode strongly resembles a dielectric waveguide mode, the coupling between both can be high, e.g. for w=0.5 µm, the power transmission is as high as 97.0% for the best cut position, this is the position inside the photonic crystal unit cell where the crystal is terminated. For wider values of w, the coupling between PhC ground

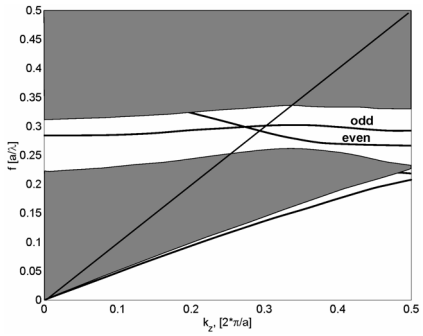


Figure 4.27: Band diagram of the triangular photonic crystal W1 waveguide with $r = 0.3a$ and $n = 2.7$.

mode and dielectric ground mode decreases but the total transmission to guided modes in the dielectric waveguide increases. For w=2.0 µm the total power transmission to guided modes is around 99.3% for the best cut position, with only the three even modes excited for symmetry reasons. When looking at figure 4.26 this means that when the PhC ground mode reaches the interface, three right-traveling guided modes are excited in the dielectric waveguide, each with a different amplitude and phase (represented by a complex excitation value). The problem is now reduced to constructing a coupler that converts this multimodal excitation as well as possible into a monomodal excitation of the broad dielectric waveguide. The advantage of this approach is that after the calculation of the PhC/dielectric scattering matrix and the optimization of the cut position, no further (time-consuming) PhC calculations are necessary and the coupler can be calculated separately from the PhC.

As a proof of principle a coupler is optimized made out of 16 sections with each a length of 0.1 µm and a width between 0.5 and 2.0 µm in steps of 12.5 nm. The first section has a fixed width, equal to the outgoing waveguide, which leaves 15 parameters to optimize. As in the optimized structure the two last coupler sections have the same width as the outgoing waveguide, these are omitted, resulting in a coupler with a total length of 1.4 µm and a transmission efficiency of 98.9%. This value is confirmed in a calculation where a semi-infinite photonic crystal waveguide excites the coupler with the ground Bloch mode. Another useful test of the optimized structure is the coupling from the broad dielectric waveguide to a finite piece of photonic crystal waveguide and back to the broad waveguide as in figure 4.28(a). The total transmission of this coupling remains above 98% as soon as there are

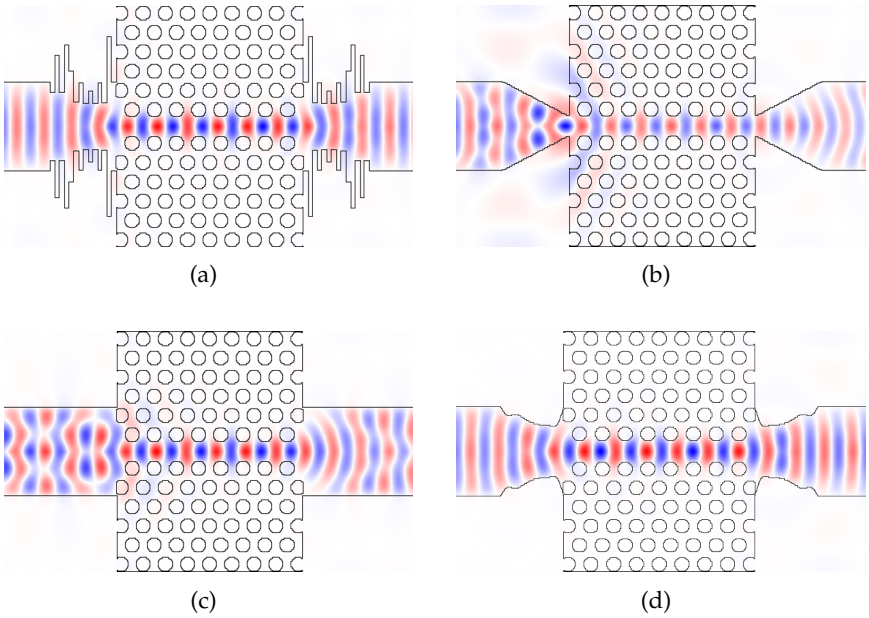


Figure 4.28: Field plot (real part of H_y) of some structures when excited from the left.

more than 3 full periods in the photonic crystal waveguide. As a comparison a linear taper (from 0.5 to 2.0 µm) with the same length as our coupler has been added in figure 4.28(b) with a total transmission of only 57% and just as an illustration the butt-coupling of the broad dielectric waveguide to the PhC waveguide is also shown in figure 4.28(c) and has a total transmission of only 41%, for an optimized cut position. In both figures interference effects from reflections are visible in the input waveguide and the curving of the field in the output waveguide, both indicating a bad conversion efficiency. A spline taper has also been optimized for the same coupling problem with a resulting efficiency of 98.0%, see figure 4.28(d).

The coupler structure of figure 4.28(a) shows very small wing-like features, in which the magnetic field has a small value. Although their exact role has not yet been determined, some are necessary for the structure to function well. When the three middle wing-like waveguide sections are replaced by sections with a width that is the average width of the two neighboring sections, the overall transmission efficiency of the structure drops below 80%. The role of the longer features is less clear and is discussed in the next chapter.

The same approach to construct a coupler between a PhC and a dielectric waveguide has been tried for other other hole diameters where the group velocity of the PhC waveguide mode is very different. In particular we have used two structures where the group velocity is half (k_z=0.364, a=0.425) and a quarter (k_z=0.389 and a=0.4205) of the above group velocity. However, we have only achieved total transmission (broad - PhC - broad) values of, respectively, 68% and 27%. The slower the group velocity of a Bloch mode, the more backwards propagating modes it will contain. In the limit of zero group velocity at the point where k_z=0.5, the Bloch mode is a standing wave, with the power evenly distributed between forward and backward waves. Our coupler is only able to excite the forward modes of the Bloch mode, which explains the strongly decreasing efficiencies when the group velocity of the Bloch mode decreases.

The content of this section is the result of a collaboration with Pablo Sanchis from the *Universidad Politécnica de Valencia* in Spain and has been mentioned in one journal and two conferences publications, see paragraph 1.6.

4.8 Conclusions

In this chapter three different methods (modified parabolic taper, block and spline taper) to construct a taper have been described and all three are successful in reducing the minimum length at which there still is a good transmission efficiency. In particular this minimum length can be half of the shortest well-performing parabolic taper.

However, this reduced taper length comes at a price. Not only does the bandwidth become smaller for shorter lengths, see figure 4.14(b) but an optimized taper shape is also only useful at the length for which it is optimized. To illustrate this some of the spline tapers of figure 4.14(a) have been recalculated using the same spline point distribution but for varying total lengths, effectively stretching and compressing them. The resulting transmissions as function of total taper length are shown in figure 4.29, where it is clear for which length each taper has been originally optimized.

Although the three construction methods clearly show some good results, the block or waveguide section approach is not suited for coupling situations where the difference in waveguide width is huge, like the studied 0.5 to 12.0 µm problem. The number of sections required for near-perfect transmission in such a situation becomes too large to be

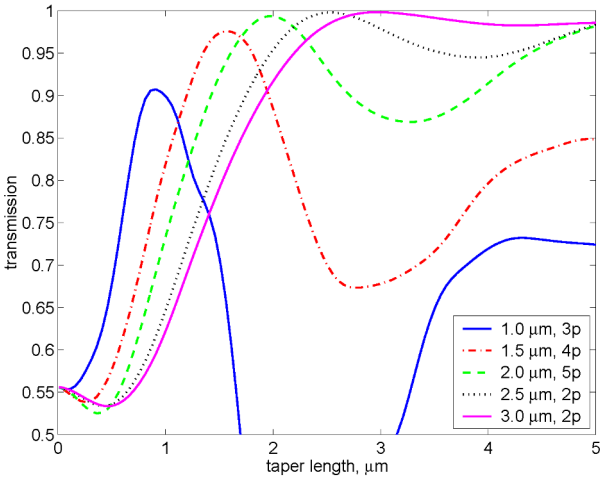


Figure 4.29: Transmission efficiency as a function of length for various spline tapers. In the legend the length for which the taper is optimized and its number of spline points is indicated.

optimized, and in this case the spline approach is the method to work with.

Most of the optimized results shown in this chapter are achieved using two-dimensional calculations after an effective index transformation. This is a good approach to verify various methods and optimization algorithms but is not accurate enough to generate structures to be implemented in a real component, as indicated in figure 4.20. As a first step in the three-directional work a unidirectional eigenmode expansion is applied for some coupling situations and good results are obtained, but when checked with a correct three-dimensional method, the results do not completely agree. This leads us, unfortunately, to the conclusion that only a bidirectional 3-D eigenmode expansion is suited to calculate real structures in a real optimization. However, an initial optimization in 2-D can still be useful as a base from which the 3-D optimization can start.

In this chapter we have also shown that the storage of some quantities during an optimization can speed up this optimization considerably, particularly in combination with the spline approach and the right discretization where the number of possible waveguide section interfaces is strongly reduced.

For further work it might also be interesting to remember that a discretization of 25 nm (in a staircase approximation) is a maximum to

make the structure appear optically smooth. This probably is dependent on the parameters used (the core and cladding refractive index and the widths of incoming waveguides) and is not be confused with sidewall roughness, where 25 nm would be detrimental.

The block approach has also been applied in other coupling problems, not between two different groundmodes, but for the conversion of more complex mode excitations. Also for these situations good results have been achieved.

Chapter 5

Interpretation

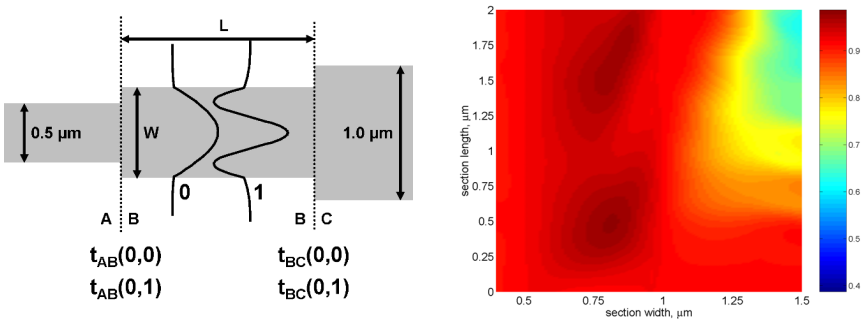
In the previous chapter various methods to construct a taper and various optimization algorithms have been discussed. These led to some promising results, indicating that it should be possible to halve the minimum taper length, when compared with the shortest well-performing parabolic taper.

The functioning of a parabolic taper is well understood, see appendix A, and it would be nice to gain the same understanding in the optimized tapers of the previous chapter, which could bring many benefits. In the most optimistic case, where we fully understand what is going on in an optimized taper, numerical optimizations might no longer be required in the future. Or somewhat less useful, a partial understanding of the operation might make it possible to recycle optimized designs to be applied in similar problems. Even if we only understand a few aspects of the optimized designs, studying them can offer additional constraints for the optimization algorithms, which can limit the parameter spaces and make the optimization faster.

In our quest for insight in the structures of the previous chapter, we will first look into the block tapers, afterwards into the spline tapers and finally try to combine the two findings.

5.1 Block Tapers

When looking at the field plots of some block tapers with a lot of sections, like figures 4.7(d) and 4.10(d), the distribution of the waveguide widths seems rather random, which of course is not the case. Initially we have tried to gain understanding in these structures by plotting the



(a) Taper comprising 1 block (b) Transmission for (W,L) sweep

Figure 5.1: Left: structure under study. **Right:** Transmission values when both the length and the width of the only waveguide section making up the taper are changed.

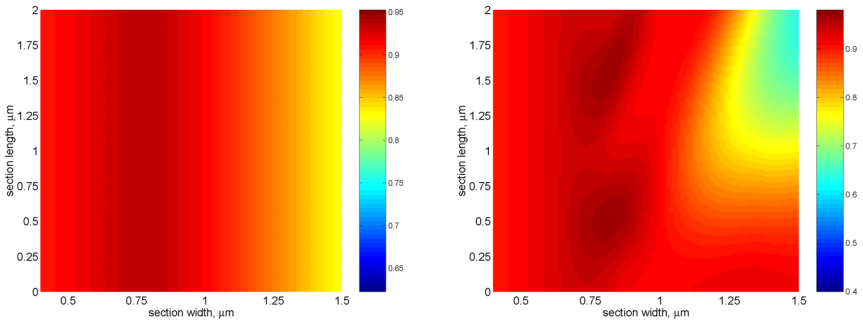
excitation of different modes as a function of z, by looking into the distribution of energy between forward and backward or guiding and radiating modes as a function of z, but all these attempts have taught us very little.

In the next pages, we try to understand block tapers by looking into the simplest cases thinkable, a block taper with just one block, then two blocks, and then some generalization into more blocks.

5.1.1 One Block

To study a taper with only 1 block by means of two-dimensional calculations, waveguides with widths of 0.5 and 1.0 µm are chosen as input and output waveguide with a core index of 2.8305 and air as cladding, see figure 5.1(a). The width W of the single waveguide section is swept between 0.4 and 1.5 µm and the length L between 0 and 2.0 µm and for each taper the ground mode transmission is calculated using CAMFR 2-D. The results of this calculation are shown in figure 5.1(b) where two islands of good performance (up to 98.7%) appear around W=0.8 and L=0.5 and 1.5 µm.

The width value seems logical, as one would expect it to lie between the two fixed widths, the periodic behavior (it repeats for longer lengths outside the plot) in the section length is somewhat less clear. To study the underlying mechanism of this periodicity, simpler calculations than a bidirectional eigenmode expansion with both guided and radiating modes have to performed. CAMFR is now only used to calculate the



(a) Forward ground mode (b) Forward first two modes

Figure 5.2: Transmission plots for the one block taper, calculated with simplified expressions containing 1 mode (**left**) and 2 modes (**right**).

transmission scattering matrix elements between the three slabs, see again figure 5.1(a). As a first, it is tested whether the forward propagation of only the ground mode can explain the values in figure 5.1(b) by simplifying the total energy ground mode transmission to:

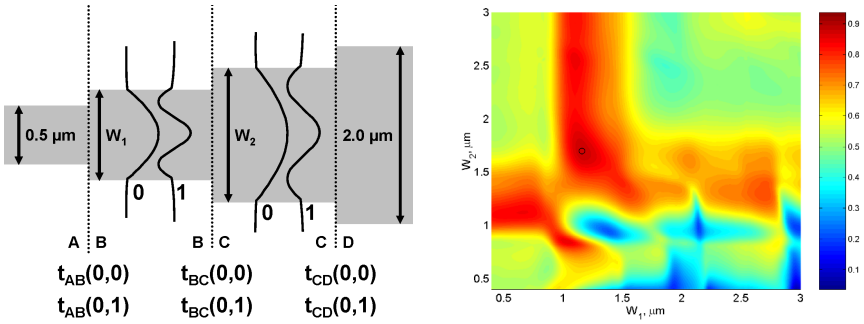
$$T = |t_{AB}(0,0) \cdot \exp(-j \cdot k_B(0) \cdot L) \cdot t_{BC}(0,0)|^2 \tag{5.1}$$

with $k_B(0)$ the propagation constant of mode 0 in slab B. This formula assumes that there is no reflection in the structure and that higher order modes never get excited. The resulting transmission for the same sweep of the structural parameters is shown in figure 5.2(a) and is length independent, what is to be expected as the phase factor containing the section length has no effect due to the modulus operation.

A next step is to take the first two forward (even) modes into account which gives as formula for the transmission:

$$\begin{aligned} T &= |t(0|0|0) + t(0|1|0)|^2 \\ t(0|0|0) &= t_{AB}(0,0) \cdot \exp(-j \cdot k_B(0) \cdot L) \cdot t_{BC}(0,0) \\ t(0|1|0) &= t_{AB}(0,1) \cdot \exp(-j \cdot k_B(1) \cdot L) \cdot t_{BC}(1,0) \end{aligned} \tag{5.2}$$

The resulting transmission, as shown in figure 5.2(b), strongly resembles the exact transmission. This proves that the periodic behavior of this taper is the result of the forward mode beating between the first two even modes, which also raises the maximum transmission from 95% in the ground-mode-only case to 98% in the exact case. Simply explained, the width of the middle section must approximately be the average of the two fixed widths to provide good ground mode transmission. The first mode that is always excited at the first interface due



(a) Taper comprising 2 blocks (b) Transmission for (W_1,W_2) sweep

Figure 5.3: Left: structure under study. **Right:** Transmission values when the widths of both waveguide sections making up the taper are changed, for constant length.

to the width mismatch can be partially be reconverted into the ground mode by appropriately choosing the length of the middle section.

5.1.2 Two Blocks

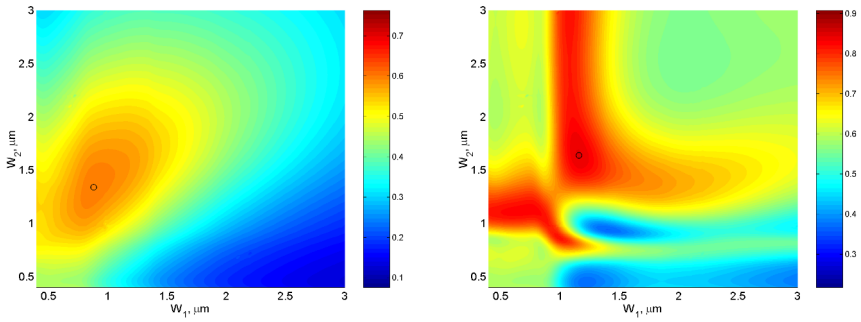
As next step a taper having two variable sections is studied. Changing four parameters simultaneously, two lengths and two widths, drastically increases the calculation time and visualizing the results at once is not possible. Therefore the lengths of the sections are chosen fixed at 1 µm and only the two widths W_1 and W_2 are varied, see figure 5.3(a).

These two widths are swept between 0.4 and 3.0 µm and the resulting correctly calculated taper transmission is shown in figure 5.3(b). The maximal transmission of 93.6% appears for W_1=1.16 µm and W_2= 1.70 and is indicated in the figure.

Using the forward propagation of only the ground mode gives as expression for the total transmission:

$$\begin{aligned} T &= |t(0|0|0|0)|^2 \\ t(0|0|0|0) &= t_{AB}(0,0) \cdot e^{-jk_B(0)L_B} \cdot t_{BC}(0,0) \cdot e^{-jk_C(0)L_C} \cdot t_{CD}(0,0) \end{aligned} \tag{5.3}$$

This formula yields the transmission plot in figure 5.4(a) with a maximum of 76.3% occurring at W_1=0.88 and W_2=1.34 µm. Both the location and the value of this maximum are well away from the correct values, the plots also look totally different.



(a) Forward ground mode (b) Forward first two modes

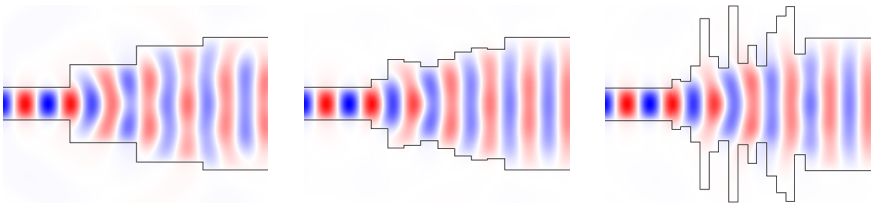
Figure 5.4: Transmission plots for the taper with two blocks, calculated with simplified expressions containing 1 mode (**left**) and 2 modes (**right**).

Adding the first mode in the equation makes it a bit more complex as there now are 4 possibilities to pass from one mode to another:

$$
\begin{aligned}
T &= |t(0|0|0|0) + t(0|1|0|0) + t(0|0|1|0) + t(0|1|1|0)|^2 \\
t(0|0|0|0) &= t_{AB}(0,0) \cdot e^{-jk_B(0)L_B} \cdot t_{BC}(0,0) \cdot e^{-jk_C(0)L_C} \cdot t_{CD}(0,0) \\
t(0|1|0|0) &= t_{AB}(0,1) \cdot e^{-jk_B(1)L_B} \cdot t_{BC}(1,0) \cdot e^{-jk_C(0)L_C} \cdot t_{CD}(0,0) \\
t(0|0|1|0) &= t_{AB}(0,0) \cdot e^{-jk_B(0)L_B} \cdot t_{BC}(0,1) \cdot e^{-jk_C(1)L_C} \cdot t_{CD}(1,0) \\
t(0|1|1|0) &= t_{AB}(0,1) \cdot e^{-jk_B(1)L_B} \cdot t_{BC}(1,1) \cdot e^{-jk_C(1)L_C} \cdot t_{CD}(1,0)
\end{aligned}
\tag{5.4}
$$

These calculations lead to the transmission plot of figure 5.4(b) where the maximum is now 90.8% and appears at W_1=1.16 and W_2=1.64, both already much better approximations of the correct values. When comparing the transmission plots of figures 5.3(b) and 5.4(b), one notices that these two correspond well. The differences that still exist are due to higher order modes and reflection within the structure. The above calculations have also been extended for 2 and more modes but the expressions become very large and the gain achieved by adding an additional mode becomes smaller.

From the quantitative studies of both the one- and the two-block taper we can conclude that their behavior can almost completely be described by taking only two modes into account, and only the forward propagation of these modes. This means that reflections only play a minor role, although we have seen in section 4.4.3 that in order to achieve the last extra percents in performance they should be taken into account.



(a) 2 sections (b) 8 sections (c) 14 sections

Figure 5.5: Field plots of block tapers with a length of 2.0 µm for coupling between 0.5 and 2.0 µm for different numbers of sections.

5.1.3 More Blocks

For tapers with more than two sections we have not conducted any quantitative analysis, but we will try to extend the previous analysis in a qualitative manner.

To this purpose the field plots of three of the block tapers of the 0.5→2.0 µm coupling situation are shown in figure 5.5, in order of increasing section count. The left plot corresponds to the maximum in figure 5.3(b) with the sections widths more or less evenly distributed between the widths of the input and output waveguides. In the center plot the taper, with 8 sections, still resembles a linear taper, but the diffraction while propagating in just one section becomes very small, meaning that some sections can deviate from the linear design without having a substantial effect on the propagation of the ground mode. These deviations serve an interference purpose for the higher order modes as in the one- and two-block tapers. In the right plot, with 14 sections, the structure no longer resembles a linear taper. The irregular width distribution makes a multi-mode interferometer of the entire structure and achieves a high efficiency, but we have not succeeded in predicting which section should have which width to reach a good efficiency. However, the mechanism that will be studied in the next section and makes the complex spline tapers perform well, is probably also responsible for the good transmission in the right plot.

Note that in the left plot, it is clearly visible that both the ground and the first even mode are excited in the middle sections and that both interfere in the right way to obtain a flat phase front in the output waveguide.

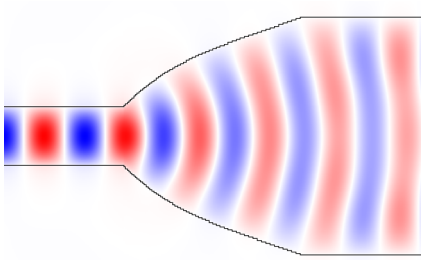


Figure 5.6: Parabolic taper with 60 sections and a length of 2 µm, input and output widths are 0.5 and 2.0 µm.

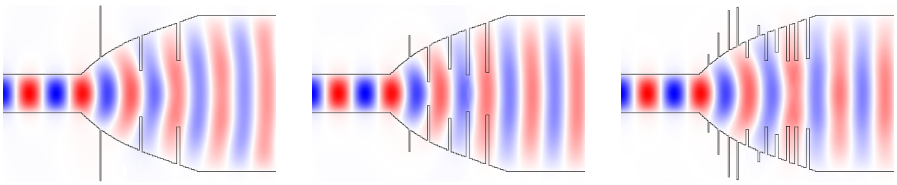
5.2 Spline Tapers

The above analysis of the simplest block tapers, having only one or two sections, has no analogue for spline tapers. As spline tapers are continuous (neglecting the discretization) by design, the coupling between modes also happens in a continuous way, and is impossible to pinpoint to one location, like the two interfaces in the one-block case. This makes a quantitative study of these tapers difficult. However, by comparing various field plots, we try to gain some general insights in these structures.

5.2.1 From 0.5 to 2.0 µm

For this coupling situation we will call tapers long if their length is at least 2.5 µm. These long tapers show no strange shapes, but mostly a monotonously increasing waveguide width between the narrow and the wide waveguide as in figure 4.15(g). These shapes deviate from a parabola but achieve a higher efficiency than the parabola for the length they are designed for. As the phase front is curved in the narrow part of these tapers, they are not adiabatic, but due to their shape manage to recover the power that initially is lost in the ground mode, as already proposed in [60] probably by a continuous version of the above forward mode beating. In tapers of these lengths, all the power seems to reside within the core throughout the structures.

The short tapers (≤ 2.0 µm) of this coupling problem show more interesting features, e.g. 4.15(h), notably some 'bumps' in the taper shape. These bumps appear in all of the short structures with enough spline points to support them and result in regions where the taper core is rather narrow. Or otherwise formulated, when the parabolic taper with

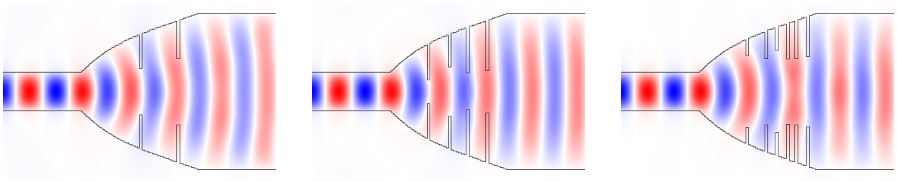


(a) 3 sections (b) 5 sections (c) 11 sections

Figure 5.7: Field plots of parabolic tapers with a number of sections optimized for better transmission. Only the section width is changed, with all lengths constant.

the same length would be drawn on top of the structures, a lot of cladding would lie within the parabola. Together with the observation that in the 'valleys' of the bumps more light is outside of the core than usual, this leads us to propose that in the short taper shapes more cladding is placed in the light path, or at least in the outer edges of the beam.

In figure 5.6 a parabolic taper with a length of 2 µm is shown with its field plot. The transmission of this taper is only 82.3% but it is interesting to see in the field plot that the losses are only due to the curved wave front, resulting in a mode mismatch. We have slightly altered this parabola to test the assumption that the transmission can be increased by carefully forcing the light to pass through some regions with a cladding refractive index. The widths of three waveguide sections are optimized by a genetic algorithm with all remaining parameters (the lengths of all sections and the widths of the other 57 sections) unchanged. This easy optimization boosts the transmission to 93.4% with the three optimized sections at z=0.5,1.0,1.5 µm, the resulting field plot is in figure 5.7(a). The same is repeated but now 5 and 11 section widths are optimized. The position of the varied sections is at, respectively, $z = i/4$ with $i = 2, \cdots, 6$ and $z = i/6$ with $i = 1, \cdots, 11$ and the resulting tapers can be seen in figures 5.7(b) and 5.7(c) with efficiencies of 98.3% and 99.7%. All three plots show similar features, with the first sections getting much longer than their initial lengths and the last sections shorter. As expected these shorter sections bring more cladding into the light path and raise the efficiency by flattening the curved wave front. Inside the cladding, where the refractive index is that of air, the light travels much faster than in the core with an index of 2.83 ($v = c/n$) and by locally increasing the phase velocity of the light at the edge of the taper, the curved wave front gets flat again. This effect is similar to the workings of a convex lens where light at the edges



(a) 3 sections (b) 5 sections (c) 11 sections

Figure 5.8: Same tapers as in figure 5.7 but the sections with protruding widths have been changed to their parabolic value.

travels less distance inside the lens material (with a higher refractive index) and is therefore speeded up when compared with light passing through the center of the lens. Based on this observation the structure in figure 4.15(h) can be seen as the combination of three diffracting sections and two converging sections (lenses), which efficiently couple the light beam between the two straight waveguides.

The function of the wider waveguide sections is less clear. To test whether they are really necessary, the tapers of figure 5.7 have been recalculated with a little modification. If a modified waveguide sections protrudes out of the taper, its value is altered to its initial parabolic value. The resulting efficiencies are 93.2%, 98.3% and 99.5% for the structures where initially 3,5 and 11 sections are optimized. These transmissions are virtually identical as before the removal of the protruding sections and prove that they do not have an optical function in the taper shapes. A possible explanation for their appearance in the first place is that their lengths are so small that diffraction within a protruding section is negligible but by taking a broader width they have more chance of still generating well-performing offspring after a mutation during the course of the genetic algorithm. Based on this reasoning their appearance is a result of the optimization and not an inherent feature of the optical design. In all the optimized spline tapers, only one is based on a spline that extends wider than the width of the output waveguide, what supports the above explanation.

In chapter 4, various taper shapes show plateaus, regions where the taper width remains constant, see figures 4.17(b) and 4.15(f). We believe that these plateaus are the step between a parabolic taper and tapers with bumps. In the field plot of figure 4.19(b) this hypothesis is supported by the fact that in the plateau region more light is outside of the waveguide core than elsewhere. This figure also demonstrates that the

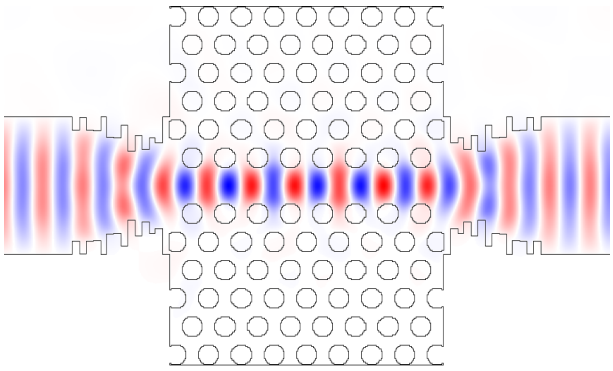


Figure 5.9: Field plot of the re-optimized photonic crystal coupler with the maximum allowed section width equal to the width of the broad waveguide.

mechanism of locally increasing the phase velocity of the light remains useful also in a rigorously 3-D calculated structure, where potentially out-of-plane losses could form a problem.

After noticing that the protruding sections in figure 5.7 do not really serve an optical goal, it seems useful to study very wide sections in other instances. The coupling situation of figure 4.28(a) is chosen to be studied again. In a first test the sections that are wider than the output waveguide are cut to the same length and the transmission is recalculated. This operation leads to a total transmission of 93.4%, which is considerably less than the 98% for the unchanged structure. However, even if the longer sections have a function in this structure, this does not necessarily mean they are necessary. To check this, the optimization of the 13 sections is run all over, with all parameters equal except that the maximal allowed width is the width of the output waveguide. The resulting structure now has a transmission of 99.2%, better than the original one and is shown in figure 5.9.

As the shape of the structure again is such that some light is forced into cladding valleys, where it propagates faster than the light in the core, we believe that the principle of this coupler is the same as in the slightly modified parabolic tapers above.

5.2.2 From 0.5 to 12.0 µm

Figure 4.17(a) in the previous chapter shows a different field plot than those for the 0.5→2.0 µm coupling situation. Practically all light remains in the core of the coupler and no small bumps appear at the

coupler's side. Striking are also the symmetrical lines of zero field strength in the center of the structure, like in a multi-mode interferometers. In the narrow part near z=5.0 μm, the ground mode is almost non-existing with most of the power in the first (even) mode. Neither the recovery of the first mode by forward mode beating nor the passing of the edge light trough cladding can explain the functioning of this structure. We believe that these structures are continuous forms of the conventional rectangular multi-mode interferometers but we have, unfortunately, not succeeded in gaining more insight in these structures. It might be possible that there is no more fundamental or more intuitive description of these structures than the rigorous calculation by eigenmode expansion.

We also believe that the bumps that appeared in the short coupling problem, do not show up here because the number of spline points studied is still too small to allow for their appearance. In order not to lose the light that is excited into radiation modes in the valley of the bump, it needs to be recaptured by guided modes quickly enough, forcing the bumps to be rather small.

5.3 Conclusion

In this chapter we have focused on some of the optimized couplers of the previous chapter and have tried to understand what is optically happening within the structure.

For the short coupling problem (0.5→2.0 μm) tapers with only 1 or 2 blocks have been studied and have revealed a mechanism where the light that is initially lost to the first even mode is recaptured using forward mode beating. In the more complex block tapers with a lot of sections and the spline tapers based on at least 3 points the role of the cladding within the light path is partially clarified. By increasing the phase velocity of the light at the edge of the coupler, curved wave fronts are more easily flattened allowing for shorter couplers as the initial curving can be compensated for. Of course there must also exist an explanation for the continuum in between the few-block tapers and tapers with the bump structure. But we have not managed to go further than the observation that interference of higher-order modes plays a role in these structures.

In the large coupling problem (0.5→12.0 μm) the exact functioning of the multi-mode interferometry still remains a mystery and we have not succeeded in identifying a model more basic than the eigenmode

expansion that we have used for the actual optimization of these structures.

In almost all studied structures, it has been found that making sections wider than the width of the output waveguide (in a narrow → wide coupling scheme) or using spline points with x-values larger than this width, are not a requirement for a well-performing optimized structure. We therefore propose to apply this observation as an extra constraint in future optimizations, opening the possibility of smaller simulation areas and faster calculations.

Chapter 6

Measurements

In order to test the designs of chapter 4 in reality, some of them have been realized in an optical chip. Our research group has a lot of experience with the common gallium arsenide (GaAs) and indium phosphide (InP) optical material systems and the newer silicon-on-insulator (SOI) material system. The latter has been chosen as the material system for the realization of this optical chip, because high resolution lithography can be applied here.

6.1 Manufacturing

Due to its excellent light confinement (refractive indices of 3.48 versus 1.0 or 1.44) and due to the maturity of silicon processing technology, silicon-on-insulator (SOI) has become an important material for research on passive photonic components. Various functional basic components have already been demonstrated in this material system, such as low-loss photonic wires and bends [62], low-loss photonic crystal waveguides [98], and high-Q cavities to couple light between different waveguides [99]. In our group active research on silicon-on-insulator involves low-loss photonic waveguides [100], vertical fiber couplers [53], arrayed waveguide gratings [63], ring resonators [61] and other optical filters [101].

For the definition of the optical functions into the SOI wafer, we make use of the CMOS processing facilities in the Flemish Interuniversity MicroElectronics Center (IMEC) in Leuven. The advantage of using techniques that have been developed for electronics and are deployed in large-scale today is that a lot of research has already been focused

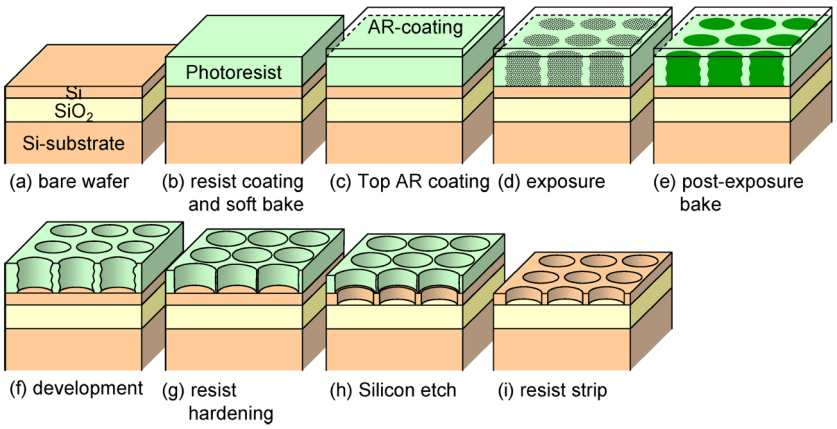


Figure 6.1: Fabrication process for photonic nanostructures in SOI using deep UV lithography and dry etching. Picture taken from [102] ©Optics Express.

on process parameter optimization and as a result the reproducibility of these techniques is excellent.

A short overview of the actual pattern definition by deep UV lithography and dry etching is shown in figure 6.1. The bare starting wafer is a disk of silicon with a diameter of 20 cm and a thickness of around 1 mm with at the top a very thin layer of silica (glass) of 1.0 µm thick and on top of that an even thinner layer of silicon of 220 nm thick. These wafers are commercially available from SOITEC and are produced using the SmartCut technology [13]. A layer of light-sensitive polymer (photo-resist) is spun on top of this wafer with a very high homogeneity and baked afterwards to get rid of the solvent used to keep the resist fluid enough. To prevent that the resist layer will form an optical cavity in the subsequent illumination step (one mirror is the air-resist interface and the other the resist-silicon interface) an anti-reflection coating is deposited on top of it.

The actual illumination or exposure is achieved by shining UV laser light with a wavelength of 248 nm through a mask that contains the features you want to project into the resist. On the mask, structures are four times larger than their planned size, which is compensated for by a 4-to-1 size reduction in the projection system. In the illuminated regions the resist is chemically altered and this process is increased by a post-exposure bake. Due to this chemical alteration, exposed resist dissolves much faster during the development step, basically leaving only the unilluminated resist behind. After an additional resist hardening

step, everything is ready to transfer the structures now present in the resist to the underlying optical layers.

Dry etching is applied for this pattern transfer, and in contrast with wet etching where the entire wafer is submerged in an acid solution, a chemically reactive ion plasma now etches into the wafer at the locations where the resist has been removed. An additional RF-field might be used to increase the energy of a plasma ion and to give it a more directional angle of incidence, thereby increasing the perpendicularity of the hole sidewalls. The vertical fiber couplers are etched to another depth (45 nm) than the rest of the structure, which requires an additional iteration of the process shown in figure 6.1. After removal of the remaining resist the pattern definition in IMEC is completed and the wafers are moved to Gent to be cleaved and characterized.

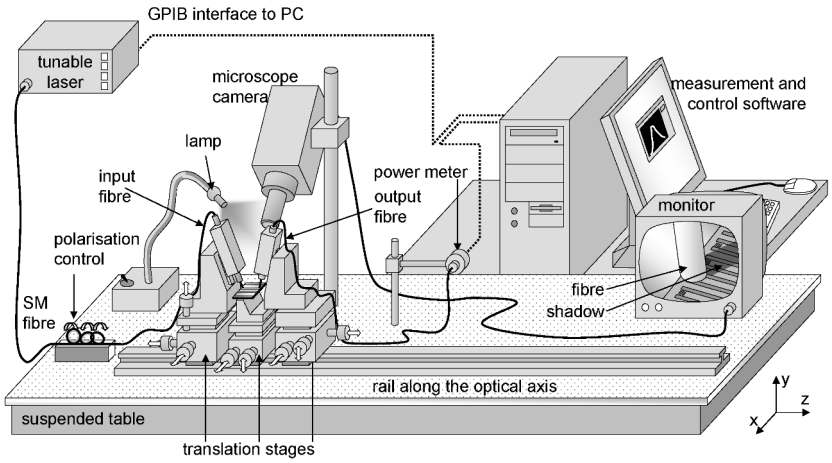
On one wafer a batch of structures, making up a die of 9x7 mm^2, is repeated several times. In one direction the same batch is just repeated identically, while in the other the dose of illumination is varied, to compensate for small variations and to make sure that at least some dies are perfectly on target.

As we have always employed vertical fiber couplers to characterize the optical components, substrate thinning and precision cleaving into millimeter-sized samples is not necessary. Instead we have manually divided the wafer into rectangles of approximately 1.5x3.0 cm^2, with at most 6 dies and 3 doses within one sample.

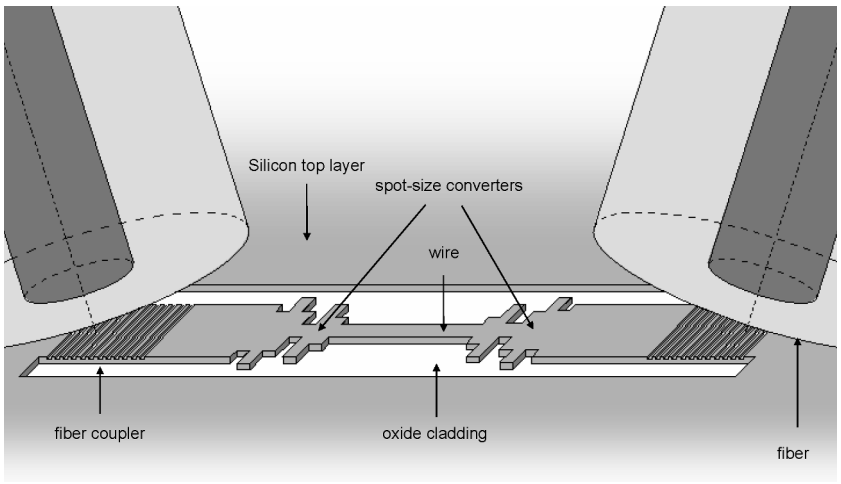
6.2 Measurement Set-up

The same measurement set-up employing vertical fiber couplers is employed for all characterizations, see figure 6.2. A tunable external cavity laser (1500-1640 nm) launches the light into a polarization-maintaining fiber. By passing through a Lefevre polarization controller [103], the polarization is turned to the desired plane, which is the electric field direction lying parallel to the grooves of the fiber coupler grating.

This fiber ends in a stripped piece that is positioned as close as possible to the fiber coupler using an x-y-z stage with manual screws assisted by piezo-electric precision screws. The fiber is inclined 10° away from the normal to the surface and lies in the plane formed by this normal and the waveguide in which the grating is defined, see figure 6.2(b). After passing through the sample under study, the light is picked up by the output fiber positioned in the same way as the input fiber. A fiber-mounted power detector registers the light output of this



(a) Detailed image of the measurement set-up



(b) Close-up on the sample and the fiber tips

Figure 6.2: Measurement set-up employing vertical fiber couplers to characterize the optical chips. Figures courtesy Wim Bogaerts.

fiber. A camera coupled to a TV screen is pointed at the sample and allows to determine which sample is being measured and it is also useful for a first visual alignment.

The close-up of a sample in figure 6.2(b) shows the fiber couplers, which consist of a one-dimensional second-order grating optimized for a wavelength of 1550 nm and for an inclination of 10° [66][67][53]. This non-vertical inclination facilitates the measurement set-up and, by breaking the symmetry, avoids reflection at the output coupler grating back into the broad SOI waveguide.

On all samples the two fiber couplers are 7 mm apart and the width of the broad waveguides in which the fiber couplers are defined is 10 µm. The structures differ in what is defined between these broad waveguides. In reference structures the broad waveguide can continue up at the output fiber coupler, or adiabatic linear tapers can connect to a short piece of photonic wire in the middle.

In the next sections various measurements are described. The name of a section corresponds to the in-house used name of the mask that has been applied to define the structures. Of the hundreds of structures that have been characterized during the course of this research only the most successful will be shown.

6.3 PICCO-03

In order to be able to quantify the transmission of the taper under study it is necessary to determine all other loss mechanisms between the tunable laser and the detector. Losses can occur due to an incomplete polarization turning, damaged fiber tips, a bad input alignment or a bad output alignment, but also within a sample not all are perfectly known. To compensate for all these losses reference structures are present in every die, which allow to normalize structures to each other.

This normalization turned out to be harder than expected. The two different reference structures present on every sample are shown in figure 6.3. In the broad reference structure, the 10 µm-wide waveguide is uninterrupted between the two vertical fiber couplers, while in the tapered reference structure the broad waveguide linearly tapers down to a photonic wire of 0.56 µm wide and 100 µm long that again tapers open to the broad waveguide. The linear tapers are 2.5 mm long, which is long enough to guarantee adiabaticity.

The spectrum of a single fiber coupler is known to closely approximate a gaussian [104], which is clearly visible for the tapered reference

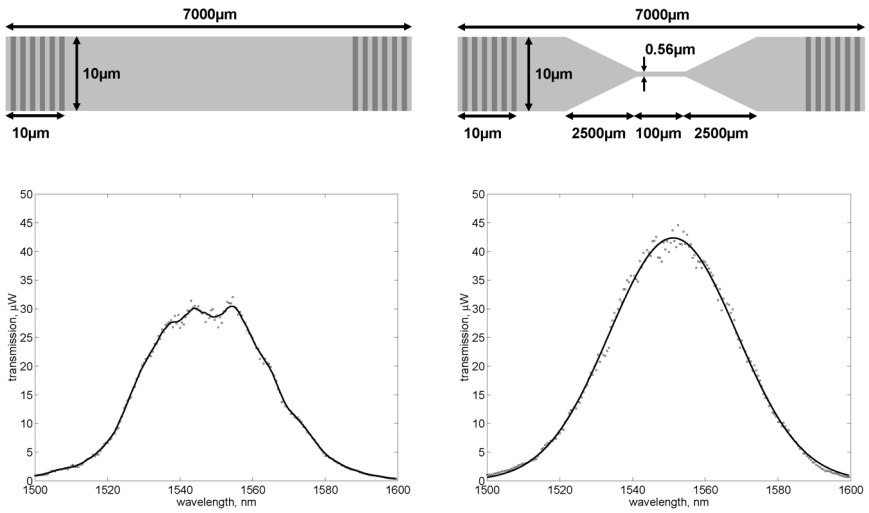


Figure 6.3: Left: broad reference structure and its measured transmission spectrum with a spectral average of these points. **Right:** tapered reference structure and its measured transmission spectrum with the best fitting gaussian curve.

structure but looks disturbed for the broad reference structure. We believe that in the broad waveguide higher order modes are somehow slightly excited, giving rise to interference features in the spectrum. The two tapers in the tapered structure act as a mode filter and reduce this interference. Assuming that the loss in the transmission of the tapered structure only arises from the fiber couplers, one calculates for an input power of 1 mW and an output power of 42.2 µW a coupling efficiency of 20.6% or -6.9 dB per fiber coupler.

Another problem with the reference structures is the spectral shift of a transmission measurement. According to which wavelength is used for the alignment of the input and output fibers, the resulting transmission spectrum can be shifted a few nm, which we will show to have a huge influence on the normalization.

Figure 6.4 shows an image of the best structure of this batch, acquired with a scanning electron microscope or SEM. It results from a block taper optimization in 2-D, with 10 waveguide sections that varied their width between 0.3 and 13.0 µm and their length between 0.3 and 3.0 µm. The resulting structure has a power transmission efficiency of 78% in a 2-D calculation, but only 71% in a rigorous 3-D calculation.

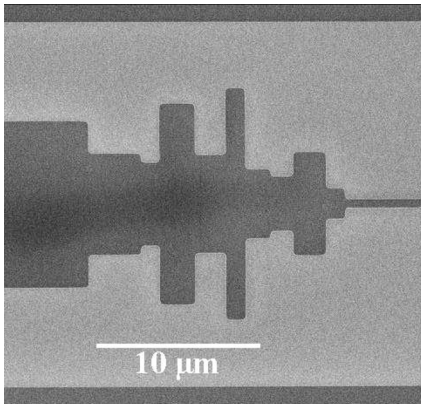


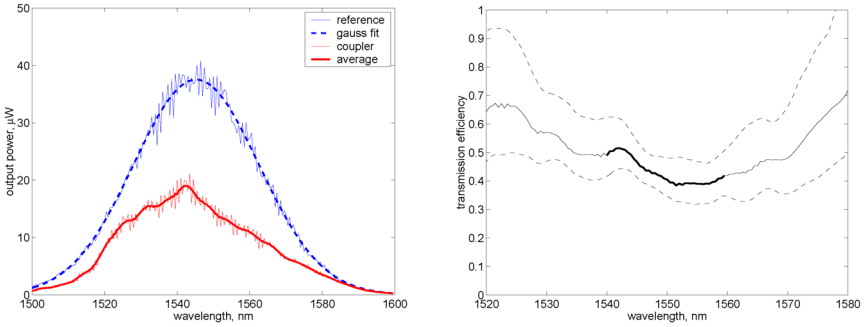
Figure 6.4: SEM picture of the component. The input and output waveguide have a width of, respectively, 10.0 and 0.56 µm. The total length of the coupler is 15.4 µm. Light gray corresponds to areas where the top silicon layer has been etched away and the underlying silica layer is revealed.

A further optimization in 3-D, using the local optimization mode of Kallistos, slightly improves the transmission up to 72%.

On the sample this structure is incorporated in a configuration as in figure 6.2(b). The transmission spectrum of the best coupler configuration is shown in figure 6.5(a) together with the transmission of the nearest taper reference structure.

A gaussian is fitted to the measured reference points and the spectral average of the measured coupler points is also shown. For a normalization of the transmission efficiency the lower curves have to be divided by the upper curves. In the low-power tails of these curves this division results in huge errors due to the noise in the measurements. But also on the slopes of the curves the normalization is not straightforward. If the reference spectrum shifts a few nm to one side, the result on the efficiency curve is huge. To quantify both effects the gaussian reference curve has been shifted 3 nm to either side and these curves are then used in the normalization. The efficiency obtained by dividing the spectral average by the gaussian fit of figure 6.5(a) is shown in figure 6.5(b) together with the upper and lower error boundaries calculated with the shifted curves.

In the region of importance (bold line), the efficiency is between 40% and 52%. As this figure includes two couplers, a single block taper has an efficiency between 63% and 72%, coming very close to the calculated maximum value for this particular structure. Measurements on similar



(a) Measured spectra (b) Normalized efficiency

Figure 6.5: Left: measured spectra of a reference structure and a block coupler with gaussian fit and spectral average. **Right:** efficiency spectrum (straight line) and lower and upper error (dashed line).

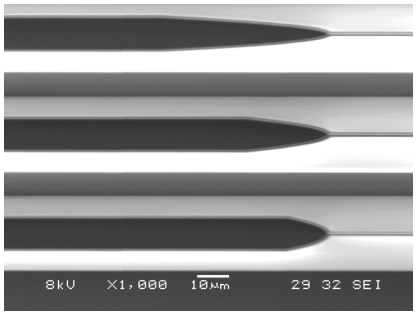
structures but with the compact block tapers replaced with short linear tapers and normalized in the same way, reveal a maximum efficiency of 40% for a single linear taper of 12.5 µm and 59% for a linear taper of 25 µm.

The above measurements, data processing and normalization have been published as a Photonics Technology Letter, see paragraph 1.6.

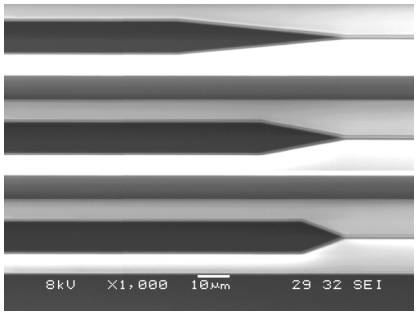
6.4 PICCO-04

A more systematic study of the performance of some block tapers in comparison with linear and parabolic tapers, is possible with the PICCO-04 structures. Figure 6.6 shows SEM images of some structures, like linear and parabolic tapers of varying length and block tapers with 5 or 10 sections. The same block taper is always repeated 7 times with a slightly different form, in an attempt to compensate for under- or over-exposure during illumination, in steps of 25 nm. This however turned out not be necessary, as the best efficiency is almost always achieved for the uncompensated structure.

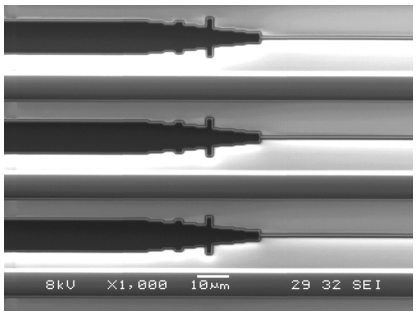
To have only a small error on the measured values, only the normalized value at maximum fiber coupling efficiency is used in the data processing of the PICCO-04 structures. This maximum lies around 1535 nm for all measured dies, and is after a first identification also used as alignment wavelength. To quantify a component's efficiency, a wavelength sweep of the output power is recorded, and a moving av-



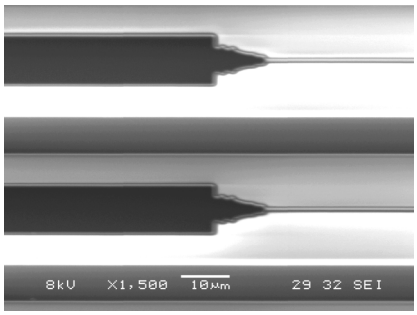
(a) Parabolic tapers



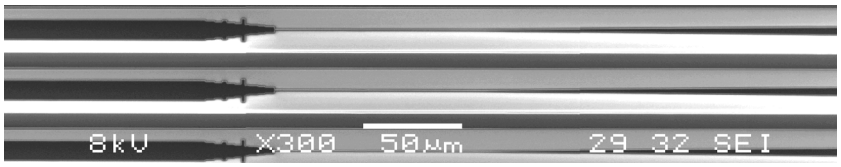
(b) Linear tapers



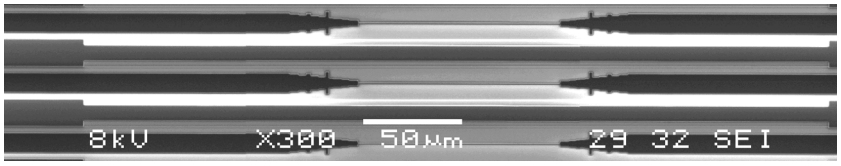
(c) Block taper with 10 sections



(d) Block taper with 5 sections



(e) Component with one block taper and one linear taper



(f) Component with two symmetric block tapers

Figure 6.6: Various components manufactured with the PICCO-04 mask. Length scale is indicated and pictures are to scale.

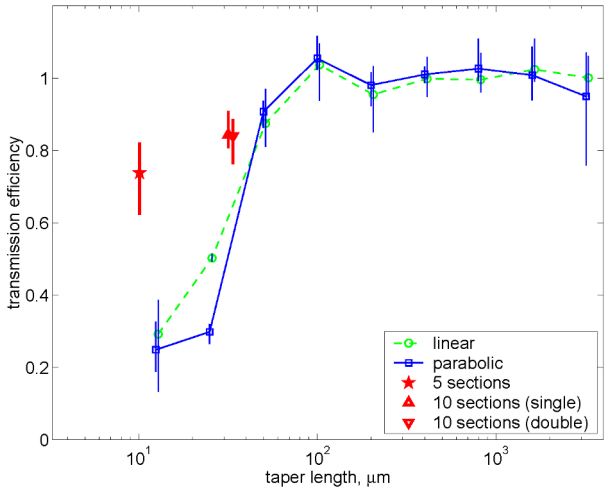
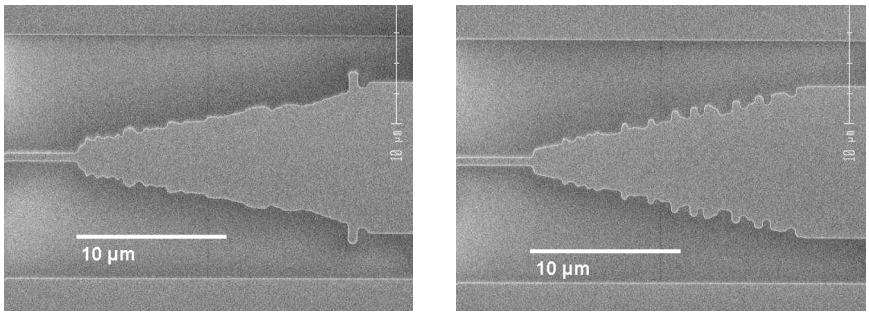


Figure 6.7: Measured efficiency values for some PICCO-04 structures. Linear and parabolic tapers with lengths of 12.5,25,50,. . .,3200 µm and block tapers with 5 (10.1 µm) and 10 (33.8 µm) sections.

erage using 9 (equally weighed) spectral points is compiled. To obtain a normalization curve, the two nearest (tapered) reference structures are measured and a gaussian is fitted to their spectrum. The envelope of these two similar gaussian curves, which closely matches these curves, is the normalization curve for all nearby structures. The spectrally averaged values are divided by this curve, and the resulting value at 1535 nm is regarded as the transmission efficiency of the total structure.

In the case of a structure with two block tapers as in figure 6.6(f), the square root of this number results in the efficiency of a single taper, while for structures as in figure 6.6(e), this extra operation is not necessary. For a single structure this efficiency measurement is repeated on several dies around the optimal exposure dose, resulting in a series of efficiency values of which the minimum, average and maximum is recorded.

Figure 6.7 combines all the PICCO-04 measurements on linear, parabolic, 5-block and 10-block tapers in one plot. The linear and parabolic tapers show a length behavior very similar to the theoretical behavior calculated in chapter 4. The efficiency of the block tapers is higher than for linear and parabolic tapers of equal length. The 5 section block taper with a length of 10.1 µm transmits 74% of the incoming light and



(a) Modified parabola (19.1 µm) (b) Modified parabola (17.8 µm)

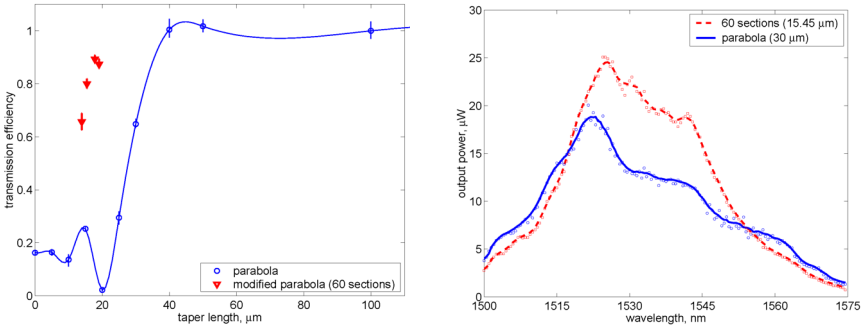
Figure 6.8: Scanning Electron Microscope images of modified parabolic tapers on the PICMOS-01 mask.

the 10-section block taper with a length of 33.8 µm up to 84%. The 3-D calculated values are respectively 70% and 79%, which lies within the error interval. It is also comforting to see that the measurements on the single and double block taper structures result in similar values.

6.5 PICMOS-01

In the last batch of structures, not only simple block tapers, but also modified parabolic tapers described in section 4.2 are characterized. SEM images of two of these tapers are shown in figure 6.8. Instead of a photonic wire width of 0.56 µm, a somewhat smaller width of 0.5 µm has been chosen to be completely sure that the second even mode in the wire (in cut-off) does not get transmitted through the short wire section between two couplers. Parabolic tapers of varying length are also present but with more lengths in the interesting region (between 0 and 200 µm) than in figure 6.7.

The results for the modified parabolas and the parabolic tapers with varying lengths can be seen in figure 6.9(a). Exactly the same process as above has been followed to normalize the measured output spectra, except for the central wavelength which now lies at 1532 nm. Because only two different dies have been characterized the error intervals are much smaller than before. Four different modified parabolas are present on the mask. In order of increasing length they are 13.9, 15.45, 17.8 and 19.07 µm long and have corresponding calculated 3-D efficiencies of 62, 69, 82 and 80%. The measured values are, for the re-



(a) Efficiencies of (modified) parabolic tapers

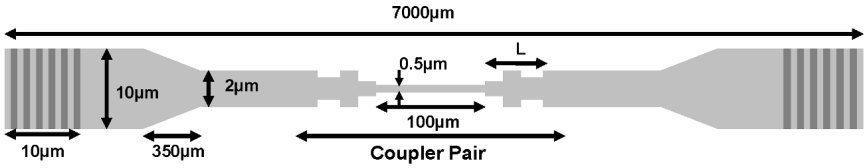
(b) Two output power spectra and the moving average of 9 points

Figure 6.9: Measurements on PICMOS-01 structures, both modified and unchanged parabolic tapers between 0.5 and 10 µm wide SOI waveguides.

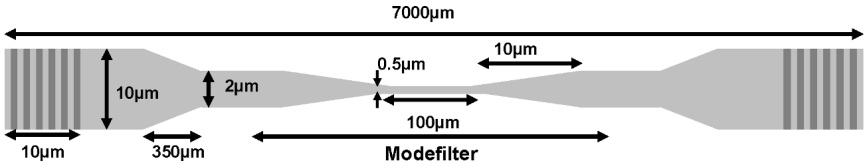
spective lengths, 66, 80, 89 and 87% and lie somewhat higher but clearly follow the same trend, see figure 6.9(a).

In figure 6.9(b) the originally measured output spectra of two of the points in figure 6.9(a) are shown. The input power for both spectra is 1 mW and as the power is measured from fiber to fiber, the light passes two couplers, and after normalization the square root is taken to obtain the efficiency of one coupler. In the region of high fiber coupler efficiency the modified parabola performs better than an unchanged parabola of almost twice the length.

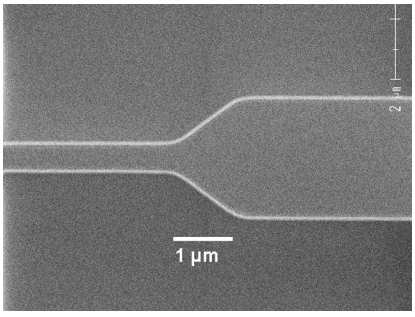
Up until now we have only studied tapers between waveguides 10 µm wide and photonic wires with a width of 0.56 or 0.5 µm. On the PICMOS-01 samples some narrower couplers are also present, namely the block tapers theoretically studied in 4.3.2, to couple between 0.5 and 2.0 µm. As the fiber couplers remain 10 µm wide, an initial adiabatic linear taper brings the waveguide width from 10 to 2 µm, see figure 6.10(a). Within the 2 µm wide waveguide the tapers under study are then located, always pair-wise. Because it is to be expected that the efficiency of a short coupler pair will be rather high, coupler pairs are also connected to each other, up to 7 pairs, making it easier to determine the efficiency of a single pair. To get rid of possible higher-order modes between various pairs, mode filters are placed in between them. A mode filter, see figure 6.10(b), is an adiabatic linear taper of 10 µm then a piece of photonic wire of 0.5 µm wide and 100 µm long and again a linear taper up to 2 µm. Block tapers containing 2 or 3 blocks and with lengths



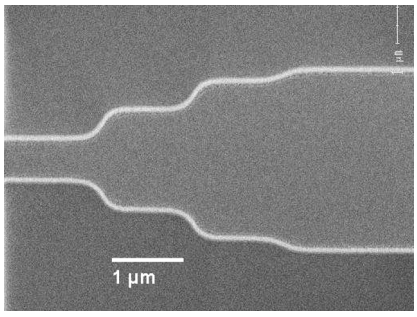
(a) Scheme of a structure with one coupler pair



(b) Scheme of a structure with one mode filter



(c) Short linear taper



(d) Two block taper

Figure 6.10: Schematics of the narrow coupler structures on PICMOS-01 and two SEM close-ups of short couplers.

of 1 and 2 µm are studied and linear tapers of the same length. SEM images of both a two-block taper and a short linear taper are shown in figure 6.10.

For every block taper the following series of configurations is available:

- N coupler pairs and $N-1$ mode filters for $N = 1, 2, 4, 8$
- $2N-1$ mode filters with $N = 1, 2, 4, 8$
- 7 coupler pairs

All output spectra are fitted with a gaussian and the value at 1532 nm is turned into a loss figure in decibels by normalization with the

Table 6.1: Measured and calculated efficiencies of short tapers.

structure	length	measured efficiency	error	calculated eff.
2 blocks	1 µm	88.5%	0.9%	93.6%
2 blocks	2 µm	94.4%	1.0%	96.4%
3 blocks	1 µm	90.5%	0.8%	93.5%
3 blocks	2 µm	95.3%	0.9%	97.9%
linear	1 µm	66.3%	9.3%	74.4%
linear	2 µm	94.0%	2.0%	91.4%

highest measured value at this wavelength. Originally we had assumed that the possible loss of a mode filter would be entirely due to the loss in the 100 µm of photonic wire, which turned out not to be the case. This leaves us with 6 unknowns to be determined, namely the efficiencies of the 4 different block tapers, the loss in a wire section and the loss of a linear taper of 10 µm long. Measurements on two dies with 9 configurations for each of the four different block tapers result in 72 loss figures. Finally this results in a system of 72 equations with 6 unknowns, for which the best fitting values are mathematically determined.

The loss of a mode filter turns out to be -0.052 dB, with -0.021 dB coming from the two linear tapers and -0.031 dB wire loss. This loss figure equals 3.1 dB/cm, which corresponds well with earlier determined values [102]. The loss values for the 4 block tapers are shown in table 6.1 together with the calculated values. Again the same trend in the figures is visible but this time the differences between theory and measurement are larger than before. The rounding of the square corners in the block tapers, see figure 6.10(d), is a possible explanation for this discrepancy.

Additionally structures are present with 7 consecutive coupler pairs, where the block tapers are replaced by short linear tapers (1 or 2 µm). Several of these structures appear identically within every die. By using the loss figure of the wire sections, the loss of a single linear taper can be determined rather easily. Table 6.1 also shows these losses with their theoretical values. For a length of 2 µm the 3-block taper performs slightly better than the linear one, at 1 µm the difference is larger and even the 2-block taper performs clearly better than a linear taper of equal length.

6.6 Conclusions

A selection of components is realized in silicon-on-insulator and characterized using vertical fiber couplers. This measurement set-up allows for a fast and simple post-processing and characterization but requires some extra caution for the data processing from measured output power to coupler efficiency. Because the smallest efficiency error appears around the central wavelength of the fiber coupler spectrum, most of the data is processed only around this wavelength.

For coupling between waveguides with widths of 0.5 (or 0.56) µm and 10 µm, tapers consisting of multiple blocks and modified parabolic tapers are characterized and compared to conventional linear and parabolic tapers. For both taper configurations, structures are identified that perform better than these conventional tapers, in the length region where adiabaticity is not yet achieved. The measured efficiencies compare well with the theoretical values.

The use of block tapers with 2 or 3 sections is studied for the coupling between waveguides with widths of 0.5 and 2.0 µm, and compared with linear tapers of equal length. Again theory and measurement correspond rather well and compact tapers perform better than there linear counterparts.

Spline tapers, as studied in chapter 4 have not been realized up until now. The study of spline tapers was performed near the end of this research work, and has come too late for incorporation in one of the described mask designs. This might, however, be a subject for further study in this area.

Although theory and experiment correspond rather well for the characterized components, no nearly lossless tapers ($<$0.1 dB) have been realized. The reason is a lack of structural optimization in a rigorous 3-D calculation scheme. Nearly al components are first optimized in 2-D, then checked in 3-D, and this 3-D value is checked against the measurements. This situation could drastically change when 3-D structural optimizations could be directly put into mask designs.

Chapter 7

Conclusions and Perspectives

7.1 Conclusions

High packaging costs and the ever increasing complexity of optical communication systems are the driving force behind optical integration and ultimately lead to photonic integrated circuits. Within these circuits various waveguide types and geometries can coexist and spot-size converters are necessary for the coupling between these waveguides. To save expensive chip area the length of these couplers should be as small as possible.

It is well-known in optical taper literature that a parabola-like taper minimizes the coupling from the ground mode to higher-order modes and is therefore the taper form that leads to the shortest adiabatic tapers. For the realization of a 3-D spot-size converter, the parabolic tapering is necessary in two dimensions, of which the vertical dimension can be technologically very hard to achieve. Various taper geometries have been proposed to tackle this problem, by using other mechanisms to broaden the beam in the vertical dimension. The application of these new broadening mechanisms in the horizontal dimension opens the way for 2-D tapers that can be shorter than conventional parabolic ones.

When designing micro-optical components, it is often useful not to rigorously calculate everything in a 3-D scheme, but to approximate an optical problem by a related 2-D problem by applying an effective index transformation. This allows for a quicker calculation time and opens the way to structurally optimize components. Genetic algo-

rithms are well suited for these kinds of problems with a high dimensionality, widely used in electromagnetic optimizations.

Three ways to construct a taper are studied in this work, one based on the modification of a parabolic taper and two ab-initio optimizations, one that directly changes the parameters of a series of rectangular waveguide sections that form the taper and one that optimizes the envelope function of these waveguide sections. In 2-D calculations all three manage to halve the minimum taper length for which a high transmission efficiency can be achieved. In 3-D these results are partially repeated using a uni-directional eigenmode expansion scheme, but it is also shown that this is a non-optimal approximation. Bidirectional 3-D calculations are still too time-consuming to repeat the optimization in a rigorous way. The suitability of the optimization methods under study is also proven for more general coupling problems, like the coupling to higher-order modes and a photonic crystal taper.

To gain understanding in the best obtained couplers after optimization, very simple structures are optimized and the optical mechanisms leading to a good coupling are identified. In simple block tapers forward mode beating between the first modes leads to a recapturing of the power that is lost at the first interfaces. In more complex structures, the alternating passing of the beam's edges through core and cladding regions, has a straightening effect on the wave fronts that get curved in the first portion of the taper. However, not all structures can be explained by these two phenomena and further study could shed some more light on the exact coupling mechanisms.

A selection of components is realized in SOI and characterized using vertical fiber couplers, which requires some caution during the data processing. The three batches that have been measured show that theory and experiment correspond rather well. In particular, in every studied coupling problem, optimized structures are found that perform better than conventional linear and/or parabolic tapers of similar length.

7.2 Perspectives

Finishing scientific research is always a somewhat artificial situation because answering a question often just raises two new ones. Also in this work, multiple problems still remain unanswered and various side paths, we have imagined could be worth looking at, remain unstudied.

More specifically, the following selection of problems could be interesting in a possible follow-up of this research work:

- The three methods for constructing a taper all have the property in common that the resulting core and cladding geometries remain separated areas. There exist no islands of cladding within the core region and vice versa. In topological terms one would say that both the core and the cladding domains are simply-connected. By allowing core and cladding islands, taper structures could become more complex and allow for higher efficiencies using the same length scales. Preliminary results from another research lab confirm this.

- Based on the observation that placing cladding material into the light path can re-bend the wave front in the right direction for a good coupling, a theoretical upper limit for transmission efficiency could maybe be calculated. On the one hand there is diffraction, which broadens a light beam upon propagation, and on the other hand the cladding regions speed the light up where necessary to re-obtain a straight wave front. For very short tapers more cladding would be necessary, but this would make the diffraction cone narrower (as it scales more or less with the refractive index). These two counteracting effects possibly lead to an optimum distribution. In a 3-D calculation, out-of-plane losses can further complicate the situation, because large regions of cladding certainly will lead to more losses.

- Neither of the two identified mechanisms can explain structures like the one in figure 4.17(a). These tapers act as generalized multi-mode interferometers and maybe a theory based on more conventional MMI's can reveal more insight.

- As already mentioned a few times, 3-D structural optimization is the only way to achieve very high efficiencies in real components. The combination of spline tapers, an intelligent discretization and the storage of already calculated scattering matrices could be the way towards this goal. The recycling of scattering matrices has in this work only been applied to unidirectional calculations, but an extension to bidirectional calculations is not unthinkable, although numerically much tougher.

- If the results of figure 4.16(b) can be repeated in a 3-D calculation, then this leads to a taper with a transmission loss of only -0.1 dB for a length of 40 µm. In combination with the theoretical fiber coupling efficiency of well below 1 dB reported by Taillaert [53],

it would become possible to realize a spot-size converter between a fiber and a wire with a footprint of only 50x15 μm^2 and a transmission loss around 1 dB. Such a component could open the way to couple 2-D fiber arrays directly onto a photonic chip, with the distance between the fiber cores only 50 µm.

Appendix A

Parabolic taper

In [23] a light ray model was used to prove for the first time that a parabola is the best shape for a regular[1] taper. Following another model, namely mode coupling, the same conclusion is drawn in [24]. Various papers have afterwards repeated this conclusion, based on different calculations methods. Two examples are [105] using the beam propagation method and the divergence angle of a waveguide mode and [106] also using the beam propagation method and mode coupling.

In this appendix we will formulate three different assumptions and work these out until the same formula 2.3. A light ray model, an eigenmode model with forward mode beating and a gaussian beam model are applied. Although this is not original work, it is added for reasons of generality.

Ray Model

Take a random taper section with width w, see figure A.1. The ground mode of a straight waveguide with this width can be written as:

$$\psi(x,z) = \cos(k_x x) \cdot \exp(-j\beta z) \tag{A.1}$$

when assuming hard waveguide boundaries, meaning that at the boundary and in the cladding the field is zero. Using the relations: $\cos(\alpha) = \cosh(j\alpha)$ and $k_x^2 + \beta^2 = k^2$ this formula can be rewritten as:

$$\psi(x,z) = \frac{\exp[-j(\cos\theta \cdot z + \sin\theta \cdot x)] + \exp[-j(\cos\theta \cdot z - \sin\theta \cdot x)]}{2} \tag{A.2}$$

[1] meaning a taper between a well-confined monomodal waveguide and a multimodal waveguide, see figure 2.1

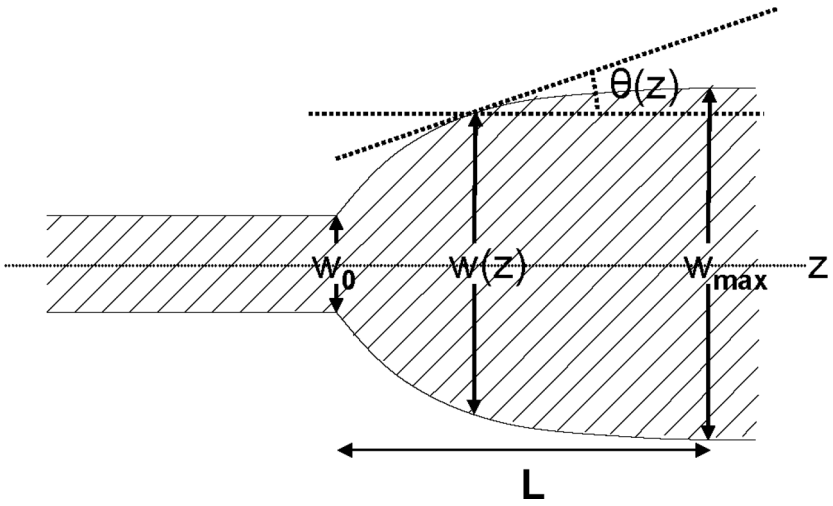


Figure A.1: A schematic parabolic taper to explain the local taper angle concept.

with $\cos\theta = \beta/k$ and $\sin\theta = k_x/k$.

This way one can think of a waveguide mode as a spatially distributed sum of rays propagating with the angles θ and $-\theta$ with respect to the propagation axis. From now on we will call this mode angle θ_m.

Assume now that a taper will be adiabatic if **the local modal angle is bigger than the local taper angle, or** $\theta_m > \theta$. Again using the hard boundary approximation the wave vector of the ground mode becomes:

$$k_{x,0} = \frac{\pi}{w} \tag{A.3}$$

and if $w >> \lambda$ the mode angle becomes:

$$\theta_m \approx \sin\theta_m = \frac{k_{x,0}}{k} = \frac{\pi}{w} \Big/ \frac{2\pi n_{core}}{\lambda} = \frac{\lambda}{2wn_{core}} \tag{A.4}$$

and the adiabatic condition becomes:

$$\theta = \frac{\alpha_1 \lambda}{2wn_{core}} \tag{A.5}$$

with α_1 a constant smaller than or equal to unity.

Gaussian Beam

Let us approximate the local taper mode by a gaussian beam with full $(1/e^2)$ beam width of $w/2$ [107]. We'll now assume that a taper is adiabatic if **the divergence angle of the local gaussian beam is bigger than the local taper angle, or** $\theta_g > \theta$.

The (half) divergence angle of a gaussian beam with (half) waist d_0 can be written as [107]:

$$\theta_g = \frac{2}{k \cdot d_0} = \frac{2 \cdot \lambda}{2 \cdot \pi \cdot n_{core} \cdot d_0} = \frac{4\lambda}{\pi \cdot n_{core} \cdot w} \tag{A.6}$$

as $d_0 = w_0/4$. The assumption now translates to:

$$\theta < \theta_g = \frac{4\lambda}{\pi \cdot n_{core} \cdot w} \tag{A.7}$$

or

$$\theta = \frac{\alpha_2 \lambda}{2 w n_{core}} \tag{A.8}$$

with α_2 smaller than or equal to $8/\pi$. A study by Hadley [105] in which the gaussian beam approximation is not used but that calculates divergence angles for various waveguides (width, index contrast and core index is varied) using BPM leads to a similar criterion.

Eigenmodes

For this paragraph, figure A.1 should be looked at from right to left, so as a taper between a broad and a narrow waveguide. In the eigenmode view, the adiabatic criterion can now be written as follows. A taper is adiabatic if **the local taper length is bigger than the local coupling length between the ground and the first even mode**. This criterion is also known as Love's strong criterion[108]. The local taper length is defined as:

$$L_{local} = \frac{w/2}{\tan\theta} \approx \frac{w}{2\theta} \tag{A.9}$$

and the coupling between the two modes is:

$$L_{coupling} = \frac{\pi}{\beta_0 - \beta_2} \tag{A.10}$$

Again using the hard boundary approximation yields for the betas:

$$\beta_i = \sqrt{k^2 - k_{x,i}^2} \approx k(1 - \frac{1}{2}\frac{k_{x,i}^2}{k^2}) = k - \frac{k_{x,i}^2}{2k} \tag{A.11}$$

$$k_{x,i} = \frac{(i+1)w}{\pi} \tag{A.12}$$

$$\beta_i \approx k - \frac{(i+1)^2\pi^2}{2kw^2} = k - \frac{(i+1)^2\pi\lambda}{4n_{core}w^2} \tag{A.13}$$

$$\beta_0 - \beta_2 = \frac{\pi\lambda}{4n_{core}w^2}(3^2 - 1^2) = \frac{2\pi\lambda}{n_{core}w^2} \tag{A.14}$$

Using the above relations the adiabatic criterion can be rewritten as:

$$L_{local} > L_{coupling} \tag{A.15}$$

$$\Leftrightarrow \frac{w}{2\theta} > \frac{\pi}{\beta_0 - \beta_2} = \frac{n_{core}w^2}{2\lambda} \tag{A.16}$$

$$\Leftrightarrow \theta < \frac{w}{2}\frac{2\lambda}{n_{core}w^2} = \frac{\lambda}{n_{core}w} \tag{A.17}$$

$$\Leftrightarrow \theta = \frac{\alpha_3\lambda}{2wn_{core}} \tag{A.18}$$

with α_3 smaller than or equal to 2.

Parabolic shape

Equations A.5, A.8 and A.18 are identical except for the numerical values of α_1, α_2 and α_3. From figure A.1 the next relation is visibly clear:

$$\theta \approx \tan\theta = \frac{1}{2}\frac{dw}{dz} \tag{A.19}$$

which combines with equations A.5, A.8 and A.18 to:

$$\frac{dw(z)}{dz} = \frac{\alpha\lambda}{w(z)n_{core}} \tag{A.20}$$

$$\Rightarrow w(z) = \sqrt{\frac{2\alpha\lambda}{n_{core}}z + C} \tag{A.21}$$

and with the right boundary condition:

$$w(z) = \sqrt{\frac{2\alpha\lambda}{n_{core}}z + {w_0}^2} \tag{A.22}$$

From equation A.20 it is easily deduced that the tangent line will be vertical if the width of the taper goes to zero, proving that the above parabola has the z-axis as symmetry axis.

Minimum taper length

Setting $\alpha = 1$, $z = L$ and $w(z) = w_L$ in formula A.22 and isolating L leads to:

$$L = \frac{\left(w_L^2 - w_0^2\right) \cdot n_{core}}{2\lambda} \tag{A.23}$$

which is a very useful relation to estimate the minimum taper length needed for adiabatic operation. Troughout this work we have called this minimum taper length the adiabatic length of a taper. Inversely, using the length and the input and output width of an existing taper, one can calculate the adiabaticity parameter with formula A.22.

Appendix B

Reciprocity and tapers

In chapter 4 tapers are mostly calculated with the narrow waveguide as input and the broad waveguide as output waveguide. In normal materials this is not a limitation as the calculation in the other direction yields exactly the same numbers. In this appendix we will clearly set the conditions for this reciprocity and apply it explicitly on the taper case. For adiabatic tapers an additional conclusion about the field distribution inside the taper is given.

General conditions for reciprocity

For monochromatic electromagnetic fields that vary harmonically in time, Maxwell's equations take the following form:

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{E} &= -j\omega \boldsymbol{B} \\ \boldsymbol{\nabla} \times \boldsymbol{H} &= \boldsymbol{J} + j\omega \boldsymbol{D} \\ \boldsymbol{\nabla} \cdot \boldsymbol{D} &= \rho \\ \boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0 \end{aligned} \tag{B.1}$$

Constitutive equations link some of these fields and in their most general form can be written like

$$\boldsymbol{J} = f_1(\boldsymbol{E}, \boldsymbol{H}) \qquad \boldsymbol{B} = f_2(\boldsymbol{E}, \boldsymbol{H}) \qquad \boldsymbol{D} = f_3(\boldsymbol{E}, \boldsymbol{H}) \tag{B.2}$$

The most general conditions for reciprocity, where $\boldsymbol{J}$, $\boldsymbol{B}$ and $\boldsymbol{D}$ may depend on both $\boldsymbol{E}$ and $\boldsymbol{H}$ can be found in [109]. If we limit these relations, however, to

$$\boldsymbol{J} = f_4(\boldsymbol{E}) \qquad \boldsymbol{B} = f_5(\boldsymbol{H}) \qquad \boldsymbol{D} = f_6(\boldsymbol{E}) \tag{B.3}$$

then reciprocity will hold under the following two conditions:

- these three functions are linear
- the resulting tensors are symmetric

Under these conditions the constitutive equations are:

$$\boldsymbol{J} = \bar{\bar{\sigma}}\boldsymbol{E} \qquad \boldsymbol{B} = \bar{\bar{\mu}}\boldsymbol{H} \qquad \boldsymbol{D} = \bar{\bar{\epsilon}}\boldsymbol{E} \tag{B.4}$$

with σ the conductive, μ the permeability and ϵ the dielectric (symmetric) tensor. The special case of isotropic media, where these tensors become scalars, is applicable in a lot of material systems.

Interpretation

The principle of reciprocity states that *if a current distribution in location A generates a electromagnetic field response in location B, then placing the same current distribution in location B will generate the same response in location A.* Note, however, that the two field distributions resulting from placing the source in the two different positions will not be identical, which can be easily understood when taking sources that emit sphere-like outward-radiating fields.

To derive Lorentz reciprocity theorem, consider two different current distributions and their resulting electromagnetic fields as in figure B.1. With the above assumptions about the interdependency, the linearity and the tensors being symmetric, Maxwell's equations can be rewritten for these fields:

$$\boldsymbol{\nabla} \times \boldsymbol{E}_1 = -j\omega\bar{\bar{\mu}}\boldsymbol{H}_1 \tag{B.5}$$
$$\boldsymbol{\nabla} \times \boldsymbol{H}_1 = \boldsymbol{J}_1 + j\omega\bar{\bar{\epsilon}}\boldsymbol{E}_1 \tag{B.6}$$
$$\boldsymbol{\nabla} \times \boldsymbol{E}_2 = -j\omega\bar{\bar{\mu}}\boldsymbol{H}_2 \tag{B.7}$$
$$\boldsymbol{\nabla} \times \boldsymbol{H}_2 = \boldsymbol{J}_2 + j\omega\bar{\bar{\epsilon}}\boldsymbol{E}_2 \tag{B.8}$$

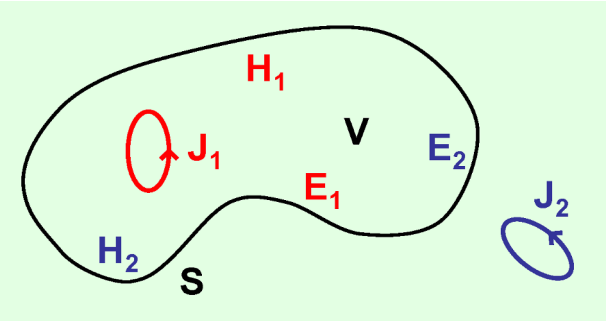


Figure B.1: Two current distributions and the resulting electromagnetic fields in and around the volume V and its contour S.

Forming the expression $\boldsymbol{H}_2 \cdot (B.5) - \boldsymbol{E}_1 \cdot (B.8) + \boldsymbol{E}_2 \cdot (B.6) - \boldsymbol{H}_1 \cdot (B.7)$, and making use of the vector identity $\boldsymbol{\nabla} \cdot (\boldsymbol{A} \times \boldsymbol{B}) = \boldsymbol{B} \cdot \boldsymbol{\nabla} \times \boldsymbol{A} - \boldsymbol{A} \cdot \boldsymbol{\nabla} \times \boldsymbol{B}$, we finally arrive at

$$\boldsymbol{\nabla} \cdot (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) = \boldsymbol{J}_1 \cdot \boldsymbol{E}_2 - \boldsymbol{J}_2 \cdot \boldsymbol{E}_1 \tag{B.9}$$

Integrating both sides over a volume V and making use of Gauss' theorem yields the following vector equation, also called the Lorentz reciprocity theorem

$$\int\int_S (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{S} = \int\int\int_V (\boldsymbol{J}_1 \cdot \boldsymbol{E}_2 - \boldsymbol{J}_2 \cdot \boldsymbol{E}_1)\, dV \tag{B.10}$$

Now choose the contour S to enclose the entire taper and the surface of the contour perpendicularly intersecting with the two straight waveguides, as in figure B.2. This divides the contour into 3 parts: a left section S_L, a right section S_R and a section S_∞ that can be chosen as far away as possible. Suppose that the volume V, enclosed by S, is current-free, which makes the RHS of equation B.10 disappear. We will assume that the fields go to zero at S_∞. This changes Lorentz relation into

$$\int\int_{S_L} (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{S} + \int\int_{S_R} (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{S} = 0 \tag{B.11}$$

Each field has at the surface a normal and a tangential part. In the above integral only the tangential part contributes. With $\boldsymbol{u}_z$ the unit vector pointing right along the optical axis, the integral can be rewritten as:

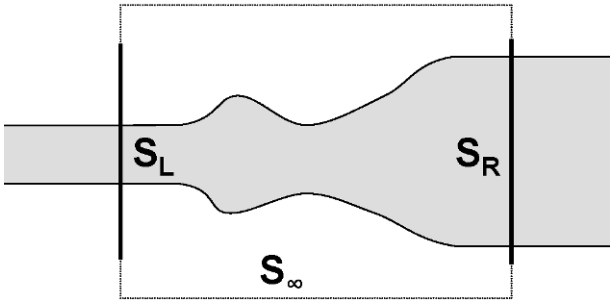


Figure B.2: The contour enclosing the taper between two straight waveguides.

$$\begin{array}{l} \int\int_{S_L} \left(\boldsymbol{E}_t^2 \times \boldsymbol{H}_t^1 - \boldsymbol{E}_t^1 \times \boldsymbol{H}_t^2\right) .\boldsymbol{u}_z dS + \\ \int\int_{S_R} \left(\boldsymbol{E}_t^1 \times \boldsymbol{H}_t^2 - \boldsymbol{E}_t^2 \times \boldsymbol{H}_t^1\right) .\boldsymbol{u}_z dS = 0 \end{array} \tag{B.12}$$

where the subscripts from above have been changed into superscripts for reasons of clarity only. The two straight waveguides are z-invariant structures and these are known to have a complete set of eigenmodes. We will assume that both waveguides have at least one guided mode and we will call the one with the highest propagation constant the ground mode of the waveguide. $(\boldsymbol{E}_1, \boldsymbol{H}_1)$ is the field distribution resulting from exciting the left waveguide with the right-traveling ground mode and analogously $(\boldsymbol{E}_2, \boldsymbol{H}_2)$ the result of exciting the right waveguide with the left-traveling ground mode. This way field $\boldsymbol{E}_1$ is known at the left surface and $\boldsymbol{H}_2$ at the right. At the opposite surface the field values depend on the exact taper shape, but due to the eigenmodes forming a complete set each field can be written as a linear combination of these eigenmodes. We know that for every eigenmode solution

$$(\boldsymbol{E}_{i,t}, \boldsymbol{E}_{i,n}, \boldsymbol{H}_{i,t}, \boldsymbol{H}_{i,n}, \beta_i) \tag{B.13}$$

there exists also a second solution corresponding to a backward propagating eigenmode:

$$(\boldsymbol{E}_{i,t}, -\boldsymbol{E}_{i,n}, -\boldsymbol{H}_{i,t}, \boldsymbol{H}_{i,n}, -\beta_i) \tag{B.14}$$

Expressing the fields at the left surface in terms of the right-propagating eigenmodes of the local waveguide:

$$\boldsymbol{E}_t^1 = \boldsymbol{E}_{0,t} \tag{B.15}$$

$$\boldsymbol{E}_t^2 = \sum_i B_i \boldsymbol{E}_{i,t} \tag{B.16}$$

$$\boldsymbol{H}_t^1 = \boldsymbol{H}_{0,t} \tag{B.17}$$

$$\boldsymbol{H}_t^2 = -\sum_i B_i \boldsymbol{H}_{i,t} \tag{B.18}$$

yields for the left surface integral:

$$\sum_i B_i \int\int_{S_L} (\boldsymbol{E}_{i,t} \times \boldsymbol{H}_{0,t} + \boldsymbol{E}_{0,t} \times \boldsymbol{H}_{i,t}) \cdot \boldsymbol{u}_z dS \tag{B.19}$$

Based on orthogonality and with the proper normalization of the modes the following relation holds:

$$\int\int_{S_L} (\boldsymbol{E}_{i,t} \times \boldsymbol{H}_{j,t}) \cdot \boldsymbol{u}_z dS = \delta_{i,j} \tag{B.20}$$

with $\delta_{i,j}$ the Kronecker delta. This simplifies expression B.19 to a form:

$$\int\int_{S_L} (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{S} = 2B_0 \tag{B.21}$$

The fields at the right surface integral and expressed in right-propagating eigenmodes of the local waveguide are:

$$\boldsymbol{E}_t^1 = \sum_i A_i \boldsymbol{E}_{i,t} \tag{B.22}$$

$$\boldsymbol{E}_t^2 = \boldsymbol{E}_{0,t} \tag{B.23}$$

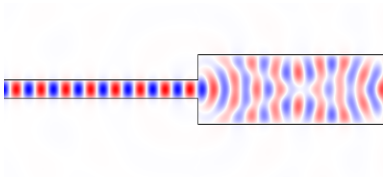
$$\boldsymbol{H}_t^1 = \sum_i A_i \boldsymbol{H}_{i,t} \tag{B.24}$$

$$\boldsymbol{H}_t^2 = -\boldsymbol{H}_{0,t} \tag{B.25}$$

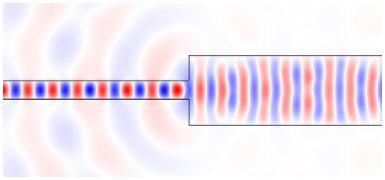
which results after an analogous calculation into:

$$\int\int_{S_R} (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{S} = -2A_0 \tag{B.26}$$

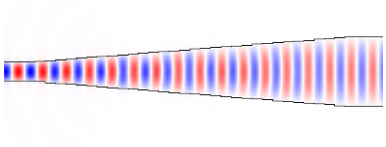
Combining equations B.11, B.21 and B.26:



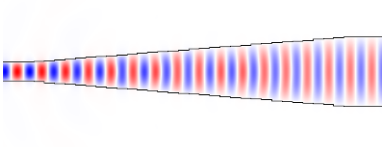
(a) Butt coupler excited from left. $t_{0,0} = -0.30 + 0.52j$



(b) Butt coupler excited from right. $t_{0,0} = -0.30 + 0.52j$



(c) Adiabatic taper excited from left.



(d) Adiabat. taper excited from right.

Figure B.3: Field distributions illustrating reciprocity and obtained with CAMFR. λ=1.55 µm, n_{core}=2.83, $n_{cladding}$=1.0, w_{in}=0.5 µm and w_{out}=1.8 µm.

$$2B_0 - 2A_0 = 0 \tag{B.27}$$

$$\Rightarrow B_0 = A_0 \tag{B.28}$$

This proves that the resulting ground mode excitation in the output waveguide will be the same after excitation from the left or from the right side.

Adiabaticity and field distributions

Consider again the situation of figure B.2 where the contour S encloses the entire taper and intersects perpendicularly with the input and output waveguides. The volume V, enclosed by S, is current-free, in which case the RHS of equation B.10 disappears. The remaining LHS only says something about the electromagnetic fields at the location of the contour. For random tapers the field distribution inside V resulting after excitation from the right (situation 1) will be different than after excitation from the left (situation 2). This changes when working with adiabatic tapers.

A taper is called adiabatic (in the broadest sense) if the ground mode transmission from input to output waveguide is perfect, so no

energy is lost to other modes. In this case it is impossible to distinguish between the field distribution resulting from the ground mode coming from the right in figure B.2 and the time-reversed field distribution resulting from the ground mode coming from the left. This demonstrates that the field inside a taper will only be identical for excitations from different sides if the taper is adiabatic. Figure B.3 illustrates this by showing both field distributions for a adiabatic and a non-adiabatic taper.

Bibliography

[1] "http://www.wikipedia/wiki/Chappe_Semaphore," *Wikipedia: an online free encyclopedia*, 2005.

[2] D. Keck, "A future full of light," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 1254–1258, November-December 2000.

[3] K. Nagayama, M. Kakui, M. Matsui, T. Saitoh, and Y. Chigusa, "Ultra-low-loss (0.1484 dB/km) pure silica core fibre and extension of transmission distance," *Electron. Lett.*, vol. 38, pp. 1168–1169, September 2002.

[4] M. Vanwolleghem, W. Van Parys, D. Van Thourhout, R. Baets, F. Lelarge, O. Gauthier-Lafaye, B. Thedrez, R. Wirix-Speetjens, and L. Lagae, "Experimental demonstration of nonreciprocal amplified spontaneous emission in a CoFe clad semiconductor optical amplifier for use as an integrated optical isolator," *Appl. Phys. Lett.*, vol. 85, pp. 3980–3982, November 2004.

[5] J. den Besten, R. Broeke, M. van Geemert, J. Binsma, E. Heinrichsdorff, T. van Dongen, E. Bente, X. Leijtens, and S. M.K., "An integrated 4 x 4-channel multiwavelength laser on InP," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 368–370, March 2003.

[6] K. Kato and Y. Tohmori, "PLC hybrid integration technology and its application to photonic components," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 4–13, January-February 2000.

[7] C. Herben, D. Maat, X. Leijtens, M. Leys, Y. Oei, and M. Smit, "Polarization independent dilated WDM cross-connect on InP," *IEEE Photon. Technol. Lett.*, vol. 11, pp. 1599–1601, December 1999.

[8] F. Xia, J. Wei, V. Menon, and S. Forrest, "Monolithic integration of a semiconductor optical amplifier and a high bandwidth p-

i-n photodiode using asymmetric twin-waveguide technology," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 452–454, March 2003.

[9] N. Kikuchi, Y. Shibata, H. Okamoto, Y. Kawaguchi, S. Oku, H. Ishii, Y. Yoshikuni, and Y. Tohmori, "Monolithically integrated 64-channel WDM wavelength-selective receiver," *Electron. Lett.*, vol. 39, pp. 312–314, February 2003.

[10] M. Castagna, S. Coffa, M. Monaco, L. Caristia, A. Messina, R. Mangano, and C. Bongiorno, "Si-based materials and devices for light emission in silicon," *Electron. Lett.*, vol. 16, pp. 547–553, March 2003.

[11] H. Rong, A. Liu, R. Jones, O. Cohen, D. Hak, R. Nicolaescu, A. Fang, and M. Paniccia, "An all-silicon Raman laser," *Nature*, vol. 433, pp. 292–294, January 2005.

[12] I. Christiaens, G. Roelkens, K. De Mesel, D. Van Thourhout, and R. Baets, "Thin-film devices fabricated with benzocyclobutene adhesive wafer waferbonding," *IEEE J. Lightwave Technol.*, vol. 23, pp. 517–523, February 2005.

[13] H. Moriceau, F. Fournel, B. Aspar, B. Bataillou, A. Beaumont, C. Morales, A. Cartier, S. Pocas, C. Lagahe, E. Jalaguier, A. Soubie, B. Biasse, N. Sousbie, S. Sartori, J. Michaud, F. Letertre, O. Rayssac, I. Cayrefourcq, C. Richtarch, N. Daval, C. Aulnette, T. Akatsu, B. Osternaud, B. Ghyselen, and C. Mazure, "New layer transfers obtained by the SmartCut process," *J. Electron. Materials*, vol. 32, pp. 829–835, August 2003.

[14] J. Zanatta, F. Noel, P. Ballet, N. Hdadach, A. Million, G. Destefanis, E. Mottin, C. Kopp, E. Picard, and E. Hadji, "Hgcdte molecular beam epitaxy material for microcavity light emitters: Application to gas detection in the 2-6 μm range," *J. Electron. Materials*, vol. 32, pp. 602–607, August 2003.

[15] G. Totschnig, M. Lackner, R. Shau, M. Ortsiefer, J. Rosskopf, M. Amann, and F. Winter, "High-speed vertical-cavity surface-emitting laser (VCSEL) absorption spectroscopy of ammonia (NH_3) near 1.54 μm," *Appl. Phys. B - Lasers and Optics*, vol. 76, pp. 603–608, May 2003.

[16] A. Gambling, "The rise and rise of optical fibers," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 1084–1093, November-December 2000.

[17] E. Yablonovitch, "Inhibited spontaneous emission in solid state physics and electronics," *Phys. Rev. Lett.*, vol. 58, p. 2059, 1987.

[18] J. Joannopolous, R. Meade, and J. Winn, *Photonic Crystals - Molding the Flow of Light*. Princeton, N.J.: Princeton University Press, 1th ed., 1995.

[19] C. Kao and G. Hockham, "Dielectric-fiber surface wave-guides for optical frequencies," *Proceedings of the IEE*, vol. 113, pp. 1151–1158, 1966.

[20] D. Marcuse, "Radiation losses of tapered dielectric waveguides," *Bell Syst. Tech. Journ.*, vol. 49, pp. 273–290, February 1970.

[21] R. Winn and J. Harris, "Coupling from multimode to single-mode linear waveguides using horn-shaped structures," *IEEE Trans. on Microwave Theory and Tech.*, vol. 23, pp. 92–123, January 1975.

[22] A. Nelson, "Coupling optical waveguides by tapers," *Appl. Opt.*, vol. 14, pp. 3012–3015, December 1975.

[23] W. Burns, A. Milton, and A. Lee, "Optical waveguide parabolic coupling horns," *Appl. Phys. Lett.*, vol. 30, pp. 3012–3015, January 1977.

[24] A. Milton and W. Burns, "Mode coupling in optical waveguide horns," *IEEE J. Quantum Electron.*, vol. 13, pp. 828–835, October 1977.

[25] B. Bransden and C. Joachain, *Introduction to Quantum Mechanics*. Harlow, UK: Longman Scientific and Technical, sixth print, 1995 ed., 1989.

[26] T. Lim, B. Garside, and J. Marton, "An analysis of optical waveguide tapers," *Appl. Phys.*, vol. 18, no. 1, pp. 53–62, 1979.

[27] N. Tzoar and R. Pascone, "Radiation in tapered waveguides," *J. Opt. Soc. Am.*, vol. 71, pp. 1107–1114, September 1981.

[28] C. Lee, L. Wu, L. Sheu, P. Fan, and J. Hsu, "Design and analysis of completely adiabatic tapered waveguides by conformal mapping," *IEEE J. Lightwave Technol.*, vol. 15, pp. 403–410, February 1997.

[29] D. Marcuse, *Theory of Dielectric Optical Waveguides*. New York: Academic Press, 1th ed. ed., 1974.

[30] A. Ash, M. Austin, and J. Love, "Efficient coupling of a laser to a waveguide using a taper designed by conformal mapping," *Opt. Quantum Electron.*, vol. 32, pp. 197–208, February 2000.

[31] S. Johnson, P. Bienstman, M. Skorobogatiy, M. Ibanescu, E. Lidorikis, and J. Joannopoulos, "Adiabatic theorem and continuous coupled-mode theory for efficient taper transitions in photonic crystals," *Phys. Rev. E*, vol. 66, p. 066608, December 2002.

[32] B. Chen, G. Tangonan, and A. Lee, "Horn structures for integrated optics," *Opt. Commun.*, vol. 20, pp. 250–252, February 1977.

[33] T. Pavlopoulos and K. Crabtree, "Optical coupling experiment with tapered (horn shaped) channel waveguides," *Opt. Commun.*, vol. 20, pp. 446–448, March 1977.

[34] J. Campbell, "Tapered waveguides for guided wave optics," *Appl. Opt.*, vol. 18, pp. 900–902, March 1979.

[35] I. Moerman, P. Van Daele, and P. Demeester, "A review on fabrication technologies for the monolithic integration of tapers with III-V semiconductor devices," *IEEE J. Select. Topics Quantum Electron.*, vol. 3, pp. 1308–1320, December 1997.

[36] J. Buus, W. Stewart, J. Willems, and R. Baets, "Spot size expansion for laser-to-fiber coupling using an integrated multimode coupler," *IEEE J. Lightwave Technol.*, vol. 11, pp. 582–588, April 1993.

[37] M. Galarza, K. De Mesel, R. Baets, A. Martinez, C. Aramburu, and M. Lopez-Amo, "Compact spot-size converters with fiber-matched antiresonant reflecting optical waveguides," *Appl. Opt.*, vol. 42, pp. 4841–4846, August 2003.

[38] K. De Mesel, R. Baets, C. Sys, S. Verstuyft, I. Moerman, and P. Van Daele, "First demonstration of 980nm oxide confined

laser with integrated spot size converter," *Electron. Lett.*, vol. 36, pp. 1028–1029, June 2000.

[39] T. Tsuchizawa, K. Yamada, H. Fukuda, T. Watanabe, J. Takahashi, M. Takahashi, T. Shoji, E. Tamechika, S. Itabashi, and H. Morita, "Microphotonics devices based on silicon microfabrication technology," *IEEE J. Select. Topics Quantum Electron.*, vol. 11, pp. 232–240, January-February 2005.

[40] K. Kawano, M. Kohtoku, M. Wada, H. Okamato, Y. Itaya, and M. Naganuma, "3-D semivectorial beam propagation analysis of a spotsize-converter-integrated laser diode in the 1.3-μm- wavelength region," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 19–21, January 1997.

[41] T. Shoji, T. Tsuchizawa, T. Watanabe, K. Yamada, and H. Morita, "Low loss mode size converter from 0.3 μm square Si wire waveguides to singlemode fibres," *Electron. Lett.*, vol. 38, pp. 1669–1670, December 2002.

[42] G. Roelkens, P. Dumon, W. Bogaerts, D. Van Thourhout, and R. Baets, "Efficient silicon-on-insulator fiber coupler fabricated using 248nm deep uv lithography," *IEEE Photon. Technol. Lett.*, To be published.

[43] K. Kasaya, O. Mitomi, M. Naganuma, Y. Kondo, and Y. Noguchi, "A simple laterally tapered waveguide for low-loss coupling to single-mode fibers," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 345–347, March 1993.

[44] O. Mitomi, K. Kasaya, and H. Miyazawa, "Design of a single-mode tapered waveguide for low-loss chip-to-fiber coupling," *IEEE J. Quantum Electron.*, vol. 30, pp. 1787–1793, August 1994.

[45] V. Almeida, R. Panepucci, and M. Lipson, "Nanotaper for compact mode conversion," *Opt. Lett.*, vol. 28, pp. 1302–1304, August 2003.

[46] Z. Weissman and A. Hardy, "2-D mode tapering via tapered channel waveguide segmentation," *Electron. Lett.*, vol. 28, pp. 1514–1516, July 1992.

[47] Z. Weissman and I. Hendel, "Analysis of periodically segmented waveguide mode expanders," *IEEE J. Lightwave Technol.*, vol. 13, pp. 2053–2057, October 1995.

[48] Z. Weissman and A. Hardy, "Modes of periodically segmented waveguides," *IEEE J. Lightwave Technol.*, vol. 11, pp. 1831–1837, November 1993.

[49] M. Spühler, D. Erni, and J. Fröhlich, "An evolutionary optimization procedure applied to the synthesis of integrated spot-size converters," *Opt. Quantum Electron.*, vol. 30, no. 5-6, pp. 305–321, 1998.

[50] Y. Rahmat-Samii and E. Michielssen, *Electromagnetic Optimization by Genetic Algorithms*. New York: John Wiley & Sons, Inc., 1999.

[51] M. Spühler, B. Offrein, G. Bona, R. Germann, I. Masserek, and D. Erni, "A very short planar silica spot-size converter using a nonperiodic segmented waveguide," *IEEE J. Lightwave Technol.*, vol. 16, pp. 1680–1685, September 1998.

[52] E. Moreno, D. Erni, C. Hafner, R. Kunz, and R. Vahldiekc, "Modeling and optimization of non-periodic grating couplers," *Opt. Quantum Electron.*, vol. 34, pp. 1051–1069, November 2002.

[53] D. Taillaert, P. Bienstman, and R. Baets, "Compact efficient broadband grating coupler for silicon-on-insulator waveguides," *Opt. Lett.*, vol. 29, pp. 2749–2751, December 2004.

[54] L. Sanchis, A. Håkansson, D. López-Zanón, J. Bravo-Abad, and J. Sánchez-Dehesa, "Integrated optical devices design by genetic algorithm," *Appl. Phys. Lett.*, vol. 84, pp. 4460–4462, May 2004.

[55] S. Kim, G. Nordin, J. Jiang, and J. Cai, "Microgenetic algorithm design of hybrid conventional waveguide and photonic crystal structures," *Opt. Express*, vol. 43, pp. 2143–2149, September 2004.

[56] J. Haes, *Study of Beam Properties of Laser Diodes and Design of integrated Beam Expansion Structures*. Ghent University, Phd. thesis, 1996.

[57] J. Haes, R. Baets, C. Weinert, M. Gravert, H. Nolting, M. Andrade, A. Leite, H. Bissessur, J. Davies, R. Ettinger, J. Ctyroky, E. Ducloux, F. Ratovelomanana, N. Vodjdani, S. Helfert,

R. Pregla, F. Wijnands, H. Hoekstra, and G. Krijnen, "A comparison between different propagative schemes for the simulation of tapered step index slab waveguides," *IEEE J. Lightwave Technol.*, vol. 14, pp. 1557–1569, June 1996.

[58] V. Kiseliov, "An optimum sectional waveguide taper for quasioptical transmission lines," *Int. J. Infrared A. Millim. Wav.*, vol. 21, pp. 163–179, February 2000.

[59] T. Felici and F. Gallagher, "On propagation through long step tapers," *Opt. Quantum Electron.*, vol. 33, pp. 399–411, April 2001.

[60] T. Felici and H. Engl, "On shape optimization of optical waveguides using inverse problem techniques," *Inverse Problems*, vol. 17, pp. 1141–1162, August 2001.

[61] P. Dumon, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. Van Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. Van Thourhout, and R. Baets, "Low-loss SOI photonic wires and ring resonators fabricated with deep UV lithography," *IEEE Photon. Technol. Lett.*, vol. 16, pp. 1328–1330, May 2004.

[62] Y. Vlasov and S. McNab, "Losses in single-mode silicon-on-insulator strip waveguides and bends," *Opt. Express*, vol. 12, pp. 1622–1631, April 2004.

[63] P. Dumon, W. Bogaerts, D. Van Thourhout, D. Taillaert, V. Wiaux, S. Beckx, J. Wouters, and R. Baets, "Wavelength-selective components in SOI photonic wires fabricated with deep UV lithography," in *Proceedings of International Conference on Group IV Photonics*, Kowloon, Hong Kong, p. WB5, IEEE-LEOS, September-October 2004.

[64] M. Smit, "New focusing and dispersive planar component based on an optical phased-array," *Electron. Lett.*, vol. 24, pp. 385–386, March 1988.

[65] Y. Barbarin, X. Leijtens, E. Bente, C. Louzao, J. Kooiman, and M. Smit, "Extremely small AWG demultiplexer fabricated on InP by using a double-etch process," *IEEE Photon. Technol. Lett.*, vol. 16, pp. 2478–2480, November 2004.

[66] D. Taillaert, W. Bogaerts, P. Bienstman, T. Krauss, P. Van Daele, I. Moerman, S. Verstuyft, K. De Mesel, and R. Baets, "An out-of-plane grating coupler for efficient butt-coupling between compact planar waveguides and single-mode fibers," *IEEE J. Quantum Electron.*, vol. 38, pp. 949–955, July 2002.

[67] D. Taillaert, H. Chong, P. Borel, L. Frandsen, R. De La Rue, and R. Baets, "A compact two-dimensional grating coupler used as a polarization splitter," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 1249–1251, September 2003.

[68] E. Miyai, M. Okano, M. Mochizuki, and S. Noda, "Analysis of coupling between two-dimensional photonic crystal waveguide and external waveguide," *Appl. Phys. Lett.*, vol. 81, pp. 3729–3731, November 2002.

[69] D. Prather, J. Murakowski, S. Shi, S. Venkataraman, A. Sharkawy, C. Chen, and D. Pustai, "High-efficiency coupling structure for a single-line-defect photonic-crystal waveguide," *Opt. Lett.*, vol. 27, pp. 1601–1603, September 2002.

[70] T. Happ, M. Kamp, and A. Forchel, "Photonic crystal tapers for ultracompact mode conversion," *Opt. Lett.*, vol. 26, pp. 1102–1104, July 2001.

[71] M. Palamaru and P. Lalanne, "Photonic crystal waveguides: Out-of-plane losses and adiabatic modal conversion," *Appl. Phys. Lett.*, vol. 78, pp. 1466–1468, March 2001.

[72] A. Talneau, P. Lalanne, M. Agio, and C. Soukoulis, "Low-reflection photonic-crystal taper for efficient coupling between guide sections of arbitrary widths," *Opt. Lett.*, vol. 27, pp. 1522–1524, September 2002.

[73] P. Sanchis, J. García, J. Martí, W. Bogaerts, P. Dumon, D. Taillaert, R. Baets, V. Wiaux, J. Wouters, and S. Beckx, "Experimental demonstration of high coupling efficiency between wide ridge waveguides and single-mode photonic crystal waveguides," *IEEE Photon. Technol. Lett.*, vol. 16, pp. 2272–2274, October 2004.

[74] A. Mekis and J. Joannopoulos, "Tapered couplers for efficient interfacing between dielectric and photonic crystal waveguides," *IEEE J. Lightwave Technol.*, vol. 19, pp. 861–865, June 2001.

[75] P. Lalanne and A. Talneau, "Modal conversion with artificial materials for photonic-crystal waveguides," *Opt. Express*, vol. 10, pp. 354–359, April 2002.

[76] P. Sanchis, J. Martí, A. García, A. Martínez, and J. Blasco, "High efficiency coupling technique for planar photonic crystal waveguides," *Electron. Lett.*, vol. 38, pp. 961–962, August 2002.

[77] P. Sanchis, J. Martí, J. Blasco, A. Martínez, and A. García, "Mode matching technique for highly efficient coupling between dielectric waveguides and planar photonic crystal circuits," *Opt. Express*, vol. 10, pp. 1391–1397, December 2002.

[78] J. Witzens, M. Hochberg, T. Baehr-Jones, and A. Scherer, "Mode matching interface for efficient coupling of light into planar photonic crystals," *Phys. Rev. E*, vol. 69, p. 046609, April 2004.

[79] Y. Xu, R. Lee, and A. Yariv, "Adiabatic coupling between conventional dielectric waveguides and waveguides with discrete translational symmetry," *Opt. Lett.*, vol. 25, pp. 755–757, May 2000.

[80] P. Bienstman, S. Assefa, S. Johnson, J. Joannopoulos, G. Petrich, and L. Kolodziejski, "Taper structures for coupling into photonic crystal slab waveguides," *J. Opt. Soc. Am. B*, vol. 20, pp. 1817–1821, September 2003.

[81] P. Barclay, K. Srinivasan, M. Borselli, and O. Painter, "Experimental demonstration of evanescent coupling from optical fibre tapers to photonic crystal waveguides," *Electron. Lett.*, vol. 39, pp. 842–844, May 2003.

[82] P. Bienstman, *Rigorous and efficient modelling of wavelength scale photonic components*. Ghent University, Phd. thesis, 2001. http://photonics.intec.ugent.be/publications/phd.asp?ID=104.

[83] P. Bienstman, H. Derudder, R. Baets, F. Olyslager, and D. De Zutter, "Analysis of cylindrical waveguide discontinuities using vectorial eigenmodes and perfectly matched layers," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 349–354, February 2001.

[84] P. Bienstman and R. Baets, "Optical modelling of photonic crystals and VCSELs using eigenmode expansion and perfectly

matched layers," *Opt. Quantum Electron.*, vol. 33, pp. 327–341, April 2001.

[85] P. Bienstman and R. Baets, "Advanced boundary conditions for eigenmode expansion models," *Opt. Quantum Electron.*, vol. 34, pp. 523–540, May-June 2002.

[86] P. Bienstman and R. Baets, "Rigorous and efficient optical VCSEL model based on vectorial eigenmode expansion and perfectly matched layers," *IEE Proceedings - Optoelectronics*, vol. 149, pp. 161–165, August 2002.

[87] P. Bienstman, "Two-stage mode finder for waveguides with a 2D cross-section," *Opt. Quantum Electron.*, vol. 36, pp. 5–14, January-February 2004.

[88] G. Hocker and W. Burns, "Mode dispersion in diffused channel waveguides by the effective index method," *Appl. Opt.*, vol. 16, pp. 113–118, January 1977.

[89] M. Qiu, "Effective index method for heterostructure-slab-waveguide-based two-dimensional photonic crystals," *Appl. Phys. Lett.*, vol. 81, pp. 1163–1165, August 2002.

[90] N. Moll and G. Bona, "Comparison of three-dimensional photonic crystal slab waveguides with two-dimensional photonic crystal waveguides: Efficient butt coupling into these photonic crystal waveguides," *J. Appl. Phys.*, vol. 93, pp. 4986–4991, May 2003.

[91] M. Dems and W. Nakwaski, "The enhanced effective index method for high-contrast photonic crystal slabs," in *Proceedings of International Workshop on Optical Waveguide Theory and Numerical Modeling*, Ghent, Belgium, OWTNM, March, 22-23 2004.

[92] W. H. Press, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge (UK) and New York: Cambridge University Press, 2nd ed., 1992.

[93] Z. Michalewicz and D. Fogel, *How to solve it : modern heuristics*. Berlin: Springer, 1998.

[94] J. Holland, "Genetic algorithms," *Scientific American*, vol. 267, pp. 44–50, July 1992.

[95] P. Vandersteegen, *Ultra-compacte koppelstructuren voor optische golfgeleiders met hoog brekingsindexcontrast*. Ghent University, master's thesis, 2003.

[96] P. Sanchis, P. Bienstman, B. Luyssaert, R. Baets, and J. Martí, "Analysis of butt coupling in photonic crystals," *IEEE J. Quantum Electron.*, vol. 40, pp. 541–550, May 2004.

[97] P. Sanchis, J. Martí, B. Luyssaert, P. Dumon, P. Bienstman, and R. Baets, "Analysis and design of efficient coupling in photonic crystal circuits," *Opt. Quantum Electron.*, To be published.

[98] M. Notomi, A. Shinya, S. Mitsugi, E. Kuramochi, and H. Ryu, "Waveguides, resonators and their coupled elements in photonic crystal slabs," *Opt. Express*, vol. 12, pp. 1551–1561, April 2004.

[99] H. Takano, Y. Akahane, T. Asano, and S. Noda, "In-plane-type channel drop filter in two-dimensional photonic crystal slabs," *Appl. Phys. Lett.*, vol. 84, pp. 2226–2228, March 2004.

[100] W. Bogaerts, R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, and D. Van Thourhout, "Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology," *IEEE J. Lightwave Technol.*, vol. 23, pp. 401–412, January 2005.

[101] P. Dumon, W. Bogaerts, D. Van Thourhout, D. Taillaert, V. Wiaux, S. Beckx, J. Wouters, and R. Baets, "Cascaded mach-zehnder filters in silicon-on-insulator photonic wires fabricated with deep UV lithography," in *Proceedings of European Conference on Optical Communications*, Stockholm, Sweden, p. Th4.2.5, ECOC, September, 5-9 2004.

[102] W. Bogaerts, D. Taillaert, B. Luyssaert, P. Dumon, J. Van Campenhout, P. Bienstman, D. Van Thourhout, R. Baets, V. Wiaux, and S. Beckx, "Basic structures for photonic integrated circuits in silicon-on-insulator," *Opt. Express*, vol. 12, pp. 1583–1591, April 2004.

[103] H. Lefevre, "Single-mode fiber fractional wave devices and polarization controllers," *Electron. Lett.*, vol. 16, pp. 778–780, September 1980.

[104] D. Taillaert, *Grating Couplers as Interface between Optical Fibres and Nanophotonic Waveguides*. Ghent University, Phd. thesis, 2001. http://photonics.intec.ugent.be/publications/phd.asp?ID=136.

[105] G. Hadley, "Design of tapered waveguides for improved output coupling," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 1068–1070, September 1993.

[106] R. Baets and P. Lagasse, "Calculation of radiation loss in integrated-optic tapers and y-junctions," *Appl. Opt.*, vol. 21, pp. 1972–1978, June 1982.

[107] B. Saleh and M. Teich, *Fundamentals of Photonics*. New York: John Wiley and Sons, 1991.

[108] J. Love, W. Henry, W. Stewart, R. Black, S. Lacroix, and F. Gonthier, "Tapered single-mode fibers and devices .1. adiabaticity criteria," *IEE Proceedings - Optoelectronics*, vol. 138, pp. 343–354, October 1991.

[109] M. Vanwolleghem, *An Integrated InP-based Optical Waveguide Isolator Using Ferromagnetic CoFe Contacts*. Ghent University, Phd. thesis, 2005. http://photonics.intec.ugent.be/publications/phd.asp?ID=145.