Een 'duobinary' ontvangerchip voor seriële communicatie aan 84 Gb/s

A Duobinary Receiver Chip for 84 Gb/s Serial Data Communication

Timothy De Keulenaer

Promotoren: prof. dr. ir. J. Bauwelinck, dr. ir. G. Torfs
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Elektrotechniek

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. D. De Zutter
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2014 - 2015

ISBN 978-90-8578-814-0
NUR 959
Wettelijk depot: D/2015/10.500/58

Universiteit Gent
Faculteit Ingenieurswetenschappen
en Architectuur
Vakgroep Informatietechnologie

Promotor: Prof. Dr. Ir. Johan Bauwelinck
Dr. Ir. Guy Torfs

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur

Vakgroep Informatietechnologie
Sint-Pietersnieuwstraat 41, B-9000 Gent, België

Tel.: +32 9 264 3340
Fax.: +32 9 264 3593

Proefschrift tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen:
Elektrotechniek
Academiejaar 2014–2015

Dankwoord

Of je het nu dankwoord, voorwoord of acknowledgement noemt, het komt allemaal een beetje op hetzelfde neer. De laatste dag voor het indienen, moet het nog geschreven worden en het is de pagina die het meest gelezen zal worden. De stress om niemand te vergeten is dan groot natuurlijk want op die bijna 5 jaar tijd is er veel gebeurd! Ik zal er dan ook voor opteren om niet alles en iedereen bij naam te noemen, je hoeft je dan ook niet ongelukkig te voelen als je hier niet met naam en toenaam vermeld wordt, ik heb alle steun en afleiding zeer zeker geapprecieerd! Onderaan heb ik trouwens een vakje voorzien voor iedereen die graag een gepersonaliseerde, gehandtekende versie wil.

In het voorjaar van 2010 was ik nog rustig samen met Arnaut aan m'n thesis bezig, en nu, voorjaar 2015, ligt de eerste ingebonden versie van m'n doctoraat hier voor mij. Dit was natuurlijk nooit gelukt mocht ik niet mogen beginnen hier in de INTEC Design groep. Hiervoor wil ik natuurlijk mijn promotoren Johan en Guy bedanken, maar ook Jan Vandewege natuurlijk die in 2010 het labo nog in goede banen leidde. Ook van Daniel, de vakgroepvoorzitter, kreeg ik altijd uitleg bij vragen die zich meer in zijn vakgebied situeerden. Daarnaast horen ook de leden van de examencommissie thuis in deze paragraaf, hen wil ik graag bedanken voor het beoordelen van mijn werk en aanreiken van interessante opmerkingen. Uiteraard mag ook het ShortTrack team niet ontbreken, bedankt voor de relevante feedback en mij op het industrieel belang van mijn onderzoek te wijzen.

Nu beginnen we aan het echte werk natuurlijk, ik ga zeker niet iedereen bij naam noemen zoals eerder al aangeven maar onder sommige namen kan ik echt niet onderuit, zoals de collega desingers van afstudeerjaar 2010: Arno, Ramses en Renato bedankt voor de ontspanning, maar ook het delen van de spanning en het samen verder bouwen aan onze grote droom: het starten van een succesvolle multinational ;)! De overige collegas en ex-collegas mogen natuurlijk ook niet vergeten worden met in het bijzonder Jean voor

de technische bijstand van het eerste uur en om de deur te openen als ik mijn batch vergeten was en Mike voor de administratieve hulp, bedankt. Van de collega's, last but by far not least natuurlijk Joris en zijn bubbel, hoewel hij hier nog maar een jaartje als collega zit en daarvoor nog een beetje een vreemd thesisstudentje was, is hij nu één van de collega's waar ik de laatste maanden het meest mee samenwerkte. Bedankt Joris, voor al het werk dat we het laatste jaar verricht hebben, als all rounder op DesignCon 2015, tijdens onze demo in Berlijn en de steeds kritische blik op de zaken, veel plezier en succes nog binnen het labo en daarna!

In het Technicum liepen natuurlijk ook nog andere mensen rond zoals de begeleiders van de Werkgroep ELEKtronica (WELEK) met onderandere Bart en Francis, de mensen van de EM groep met natuurlijk thesis partner Arnaut en regelmatige project partner Thomas in het bijzonder, de vrienden bij de Floheacom zoals ondermeer Kathleen en Bernd. Bedankt om de zorgen voor de nodige ontspanning!

Het labo en bij extensie het Technicum is bij deze overlopen, dan zijn er natuurlijk nog alle buiten labo'se activiteiten, niet dat ik zo een ongeloofelijk druk bezet persoon ben, want ik kan best genieten van een rustige avond tv kijken samen met mijn favoriete vriendin allertijden Ellen, maar sommige dingen konden we natuurlijk niet laten voorbijgaan. Om een beetje te beginnen bij het begin zitten we terug in 2010 waarbij we nog wekelijks een goed gevuld en vullend pasta team organiseerden en de zoektocht naar een huis startte! Een jaar later was het dan zover en konden Toon, Bernd, Ellen en ikzelf intrekken in de Ham 54! Merci allemaal! Tijdens het werken werd het al snel duidelijk dat ontspanning niet te onderschatten is! Bij deze: Wieland, Quinten, Stijn, Joris en de hele bende bedankt voor het mede entertainen, boeken, plannen en rijden tijdens onze zomer- en wintervakanties! Stijn en Bernd, bedankt voor het organiseren van de jaarlijkse bbqs.

Onder een apparte vermelding van Bernd kan ik natuurlijk niet onderuit als mede-burgie, mede-eigenaar van Ham 54, en zoveel meer: je stond altijd paraat met raad en daad! Bedankt!

En dan komen we natuurlijk tot bij de familie, in het begin de harde kern: de mams, de paps en de zussen (ons Julie en ons Amelie) en natuurlijk 'mijn' Ellentje. Later werd de groep dan nog wat groter en kwamen daar ook nog Günther en Sarah bij. Hen allen wil ik niet enkel bedanken voor het nalezen

en tonen van oprechte interesse, maar ook voor de steun, ontspanning onder de vorm van kubb, bbqs en molteams en het maken van mij tot hoe ik ben geworden :) De familie is natuurlijk meer dan enkel het gezinnetje De Keulenaer, daar horen ook oma en opa bij, die regelmatig eens belden om op de hoogte te blijven, of was het om een technisch probleem op te lossen? En de nonkels en tantes, bedankt!

Nu volgt natuurlijk nog de belangrijkste bedanking, die voor Ellen, ik mocht niet teveel koosnaampjes gebruiken en heb dat dan ook proberen vermijden. Bedankt voor alle steun, het nalezen, de interesse, het luisteren, het zorgen voor de plantjes en nog zoveel meer!

Gent, mei 2015
Timothy De Keulenaer

If two people always agree, one of them is useless.
If they always disagree, both are useless.

MARK TWAIN

Table of Contents

Dankwoord i

Nederlandse samenvatting xiii

English summary xvii

List of Publications xxi

1 Introduction 1

1.1 Background 1
1.1.1 Backplane configurations 2
1.1.2 Modulation formats 6
1.1.3 Duobinary system 10
1.2 Overview of the work 13
1.2.1 Chip design and analysis 13
1.2.2 Measurements and results 14
1.3 Outline of the dissertation 14

I Design and analysis **19**

2 Receiver design 21

2.1 Introduction and design overview 22
2.2 Design of a 80 Gbit/s SiGe BiCMOS fully differential input buffer for serial electrical communication 24
2.2.1 Introduction 24
2.2.2 Block diagram 24
2.2.3 Circuit description 25
2.2.3.1 Input buffer 25
2.2.3.2 Output buffer 25
2.2.4 EM modeling 25

2.2.5 Simulation results 26
2.2.5.1 Small-signal simulations 27
2.2.5.2 Time-domain simulations 27
2.2.6 Comparison 28
2.2.7 Conclusion 29
2.3 A Digitally Controlled Threshold Adjustment Circuit in a 0.13 µm SiGe BiCMOS Technology for Receiving Multi-level Signals up to 80 Gb/s 32
2.3.1 Introduction 32
2.3.2 Non-clocked duobinary receiver 32
2.3.3 Emitter follower threshold adjustment circuit . . . 33
2.3.3.1 Emitter follower threshold adjustment . 33
2.3.3.2 Cascoded output driver 36
2.3.3.3 Bandwidth and linearity 36
2.3.4 Digitally controlled current DACs 36
2.3.5 Fabrication and further Work 37
2.3.6 Conclusion and results 39
2.4 Additional high-speed subblocks 39
2.4.1 Combining the level shifters into the LSLA 39
2.4.2 Sampling stages 40
2.4.3 Half rate XOR gate 41
2.4.4 Quarter rate output buffers 42
2.4.5 Clock chain 43
2.5 Power consumption overview 44
2.6 Conclusions 45

3 Measurements of millimeter wave test structures for high-speed chip testing 49
3.1 Introduction 50
3.2 Design of millimeter wave test structures 51
3.2.1 Coupled GCPW 51
3.2.2 Coupled to uncoupled GCPW splitter 52
3.2.3 Connector footprint 53
3.2.4 Rat-race coupler 55
3.3 Measurements of high-speed test boards 56
3.3.1 Material characteristics 56
3.3.2 Connector characteristics 58
3.3.3 Bandwidth of differential lines 58
3.3.4 Rat-race coupler 59

3.4 Flip chip mounting of the receiver chip 60
3.5 Design of a flip chip test board 65
3.6 Design of a multilayer active daughter card 68
3.6.1 Board stack up 68
3.6.2 Routing 72
3.6.3 Mounting order 74
3.6.4 Board performance 75
3.7 Conclusion 77

II Measurements and results 81

4 84 Gb/s SiGe BiCMOS duobinary serial data link including Serialiser/Deserialiser (SERDES) and 5-tap FFE 83
4.1 Introduction 84
4.2 Transmitter 85
4.3 Receiver 86
4.4 Measurements and conclusion 89

5 56+ Gb/s serial transmission using duobinary signaling 93
5.1 Introduction 94
5.2 Duobinary signaling 95
5.3 Custom ASIC design for 50+ Gb/s duobinary link 100
5.3.1 Feed-forward equalizer 100
5.3.1.1 Introduction to feed-forward equalization 100
5.3.1.2 Implementation 12.4 ps spaced 5-tap FFE 100
5.3.1.3 Parameter optimization 102
5.3.2 Duobinary receiver 103
5.3.2.1 Introduction to duobinary receivers . . . 103
5.4 Eye-pattern and BER measurements 106
5.4.1 Measurement setup 106
5.4.2 Measurements on ExaMAX® demonstrator 107
5.5 Measurements on active daughter cards 111
5.6 Conclusions 114

6 Conclusions and future research 121
6.1 Summary of the results 121
6.2 Valorisation opportunities 123
6.3 Future research 124

Nederlandse samenvatting –Summary in Dutch–

De laatste jaren is er een enorme groei van het dataverkeer in de telecom-, datacom- en supercomputersector. Gebruikers verwachten hoge-kwaliteit, bandbreedte verslindende, 4K-video conferenties met meerdere gebruikers. Ze verwachten dat Google in één milliseconde door zijn honderden terabytes grote database kan zoeken en dat een supercomputer pi kan berekenen tot 10 miljoen getallen na de komma in minder dan een microseconde. Om al deze verwachtingen te kunnen inlossen, zal de communicatiesnelheid in server racks drastisch moeten opgeschaald worden. Typisch gebruiken deze racks een architectuur waarbij verschillende kaarten in een backplane geplugd worden. Dit kunnen lijnkaarten zijn voor telecom, opslagmodules voor data centers, verwerkingseenheden voor supercomputers, … Hogesnelheidscommunicatie over deze backplanes is zeer uitdagend omwille van hun zeer hoge frequentieafhankelijke verliezen. Om dit euvel op te lossen kan men afstappen van de typische on-off keying non-return-to-zero modulatie en in de plaats daarvan gebruik maken van meer bandbreedte-efficiënte modulatieformaten zoals een 4-niveaus puls amplitude gemoduleerd signaal (PAM4) of een 'duobinary' gemoduleerd signaal.

De interesse om over te stappen op andere, meer bandbreedte-efficiënte, modulatieformaten is er niet enkel voor backplane communicatie, ook bij korte-afstands optische netwerken is hier veel onderzoek naar. De INTEC Design onderzoeksgroep van de Universiteit Gent is al verschillende jaren bezig met onderzoek naar PAM4 en 'duobinary' modulatie, voor zowel elektrische als korte-afstands optische netwerken. Dit om efficiënter gebruik te maken van de kanaalbandbreedte en zo de gewenste datasnelheden te bereiken. In dit werk is gekozen voor 'duobinary' modulatie omwille van de afweging tussen bandbreedte-efficiëntie, implementatiecomplexiteit en vermogenverbruik.

Duobinary werd voorgesteld in 1963 door A. Lender in zijn paper 'The

duobinary technique for high-speed data transmission'. Hij toonde aan dat een beperkte hoeveelheid toegevoegde complexiteit kan leiden tot een veel efficiënter gebruik van de kanaalbandbreedte. Dit zou vooral van belang zijn bij bedrade en hoogfrequente radiocommunicatie. Meer dan 40 jaar later beschreef J. Sinsky dat 'duobinary' modulatie ook interessant kan zijn bij backplane communicatie. In zijn artikel toont hij een vernieuwende techniek voor het decoderen van een 'duobinary' gemoduleerd signaal tot een traditoneel digitaal signaal door gebruik te maken van een XOR poort in plaats van een gelijkrichter (zoals voorgesteld door A. Lender). Deze techniek zorgde ervoor dat 'duobinary' terug kan concurreren met PAM4 voor hogesnelheidscommunicatiekanalen. Duobinary is verschillend van de meer courante modulatieformaten omdat het gebruik maakt van intersymboolinterferentie en zo de kanaalverliezen benut in plaats van ze volledig te proberen compenseren. Duobinary is een speciale vorm van partial response signaling met een specifiek decodeeralgoritme: het 'duobinary' gemoduleerd signaal wordt gevormd aan de uitgang van het totale kanaal (aan de ingang van de ontvanger). Dit kanaal wordt op zijn beurt gevormd door een combinatie van het predistorsie filter, bv. geïmplementeerd als een feedforward equalizer, en het transmissiekanaal.

Het werk voorgesteld in deze thesis is gebaseerd op het ontwerp en het testen van een hogesnelheids-duobinary-ontvanger. Het ontwerp en de analyse van de chip en bijbehorende testborden worden grondig beschreven. De chip bestaat uit twee delen: een eerste deel waarin de 'duobinary' gemoduleerde data ontvangen wordt en een tweede deel waarin de ontvangen data gedecodeerd wordt. Verschillende onderdelen van het eerste gedeelte werden door mezelf gepresenteerd op internationale conferenties. Het belangrijkste onderdeel, de ingangsbuffer, is geïmplementeerd als een transimpedantieversterker aangepast aan een $100\,\Omega$ differentiële impedantie en werkt als een zeer efficiënte breedbandige lageruisversterker. Deze heeft een gesimuleerde bandbreedte van 50 GHz, een ruisgetal lager dan 5 dB en een versterking van 12 dB. Verder zijn er twee niveau-verschuivende begrenzende versterkers verbonden met deze ingangsbuffer. Elke versterker bestaat uit twee niveau-verschuivende elementen met enkele versterkende elementen er tussenin. Door verschillende niveau-verschuivende elementen te gebruiken, kan zowel de bandbreedte als het bereik waarover het niveau verschoven kan worden, nodig om het onderste en het bovenste oog van een 'duobinary' signaal te ontvangen, gemaximaliseerd worden. Het tweede deel van de chip bestaat uit een demultiplexer en decodeer-structuur. Eerst wordt het signaal gesplitst in vier signalen, elk aan de halve snel-

heid. Deze worden dan gedecodeerd door twee XOR poorten die werken aan halve snelheid. Hierna worden deze twee datastromen verder gesplitst tot vier datastromen aan een kwart van de snelheid, die dan naar de uitgangsbuffers gestuurd worden. De technieken die in de complete chipset geïmplementeerd zijn, zijn het onderwerp van 3 patentaanvragen.

Om systeemtesten met de ontworpen chipset mogelijk te maken, is er nood aan een testbord dat de prestaties van de chips niet beperkt. Hiervoor werden verschillende connectoren, bordmaterialen en routeertechnieken onderzocht en vergeleken. De resultaten hiervan werden door mezelf gepresenteerd op de IEEE Workshop on Signal and Power Integrity in 2014. Gebruik makend van 1.85 mm Rosenberger connectoren werden er verbindingen getoond met een vlakke verlieskarakteristiek tot 67 GHz, inclusief het veranderen van de lijnbreedte, tussen de connector en de chip voetafdruk, en de overgang van ongekoppelde naar gekoppelde transmissielijnen. Deze techniek samen met het thermosonisch flip-chip monteren van de chip op het bord, leidde tot meerdere succesvolle testborden, gaande van een klein 4 laags Rogers RO4003C bord met maar één chip op, tot een groot 24 cm x 30 cm 12 laags Megtron 6 bord met twee ontvanger- en twee zenderchips.

Het combineren van de ontworpen chip met het vier lagen testbord leidde tot een 84 Gb/s transmissielink over 20 cm coax-kabel verbonden tussen de twee testborden met minder dan één bitfout per 100 miljard bits (BER $< 10^{-11}$). Het kanaalverlies bij 42 GHz was ongeveer 14 dB en werd gevormd tot een 'duobinary' kanaal door een 5 taps feedforward equalizer aan de zendkant. Tot nu toe is dit de snelste seriële 'duobinary' gemoduleerde link. Dit was dan ook de aanleiding van een journal paper gepubliceerd in Electronic Letters.

Tot slot werd er ook een compleet backplane testplatform gemaakt, om de performantie van de chip op 56 Gb/s te demonstreren over grotere afstanden. Het testplatform bestaat uit twee borden met elk vier chips die op een Megtron 6 FCI ExaMAX® backplane demonstrator geplugged kunnen worden. Foutvrije (BER $< 10^{-13}$) communicatie werd aangetoond bij link lengtes tot 50 cm met verliezen boven de 40 dB bij 28 GHz. Een 56 Gb/s link met 33 dB verlies op 28 GHz werd live gedemonsteerd op DesignCon 2015 in Santa Clara (USA) op de stand van FCI. Dit met een 28 Gb/s crosstalk stoorbron. Samen met Jan De Geest presenteerde ik een paper hierover op DesignCon 2015 die beloond werd met een best paper award.

Het werk voorgesteld in deze thesis kreeg veel aandacht van de industrie tijdens een demo op de Bell Labs Future X dagen en op DesignCon 2015. De ontworpen chipset en bordontwerptechnieken kunnen gebruikt worden voor verschillende doeleinden, zoals: het aansturen van korte-afstands hogesnelheids optische kanalen, communiceren over complexe backplane kanalen met veel frequentieafhankelijk verlies of voor korte chip-to-module toepassingen.

Samen met drie andere onderzoekers van de INTEC Design groep hebben we recent financiering verkregen van het IOF UGent fonds om de technologie ontworpen tijdens deze thesis verder te ontwikkelen en klaar te maken voor industrieel gebruik.

English summary

During the last decades, the data traffic in telecom, datacom and high performance computers (HPC) is skyrocketing. Users expect high quality, bandwidth devouring, 4K multi-user video conferencing; they expect Google to query their hundreds of Terabytes database in milliseconds and expect supercomputers to calculate 10 million digits of Pi in a microsecond. To keep up with these expectations, the communication speed within server racks needs to increase dramatically. Typical rack systems use an architecture where multiple cards are plugged into a big backplane. These can be, among others, line cards for telecom, storage blades for datacom, blade servers for HPCs. High-speed communication across these backplanes is very challenging, due to the high amount of frequency dependent loss. To overcome these challenges, one can look at trading in the typical on-off keying non-return-to-zero communication methods for more bandwidth-efficient formats such as 4 level pulse amplitude modulated signals (PAM4) or duobinary modulated signals.

The interest in moving towards more spectrally efficient modulation schemes is not only coming from backplane communication, also in short reach optics numerous studies are ongoing regarding this topic. The INTEC Design research group of Ghent University has been investigating PAM4 and duobinary for a couple of years, for use in electrical as well as short reach optical interconnections, to make more efficient use of the total available channel bandwidth and achieve the desired increase in data rate. The selection of duobinary in this work is based on the trade off between spectral efficiency, implementation complexity and power consumption.

Duobinary was first introduced by A. Lender in 1963 in his paper 'The duobinary technique for high-speed data transmission'. He showed that a low degree of added complexity could result in a significantly more efficient use of the channel bandwidth. This would mainly be beneficial for wireline and high-frequency radio communications. More than 40 years later, J. Sinsky wrote about using duobinary signaling for electrical back-

plane systems. He described a novel technique of decoding a duobinary modulated signal to a traditional binary signal using an XOR gate instead of a full wave rectifier (as proposed by A. Lender). This technique allowed duobinary signaling to compete on record breaking rates with its main rival PAM4. Duobinary signaling differs from the other more commonly known modulation formats: it is based on intersymbol interference and in this way embraces the channel loss in contrast to fighting it. It can be thought of as a special case of partial response signaling with a specific decoding structure. A duobinary modulated signal is generated at the receiver by combining the channel loss and a predistortion filter, such as a feed forward equalizer, to shape the channel.

The work presented in this dissertation is based on the design and testing of a high-speed duobinary receiver. The design and analysis of the chip and the accompanying test boards are thoroughly discussed. The chip consists of two main parts: a duobinary front end and a DEMUX/XOR back end. Multiple blocks of the duobinary front end were published and presented at international conferences by myself. The main block of the front end is the input buffer which is implemented as a transimpedance amplifier matched to a 100 Ω differential impedance and operates as a very effective wideband low noise input buffer. In simulations, the input buffer has a bandwidth of 50 GHz, a noise figure below 5 dB and a gain of 12 dB. Subsequently, two level shifting limiting amplifiers are connected to the output of the input buffer. Each consists of two level shifting stages and different amplifying stages. Using multiple level shifting stages allows the designer to maximize the bandwidth and level shifting range necessary to convert the upper and lower duobinary eyes to traditional NRZ eyes. The back end of this chip consists of a DEMUX/XOR structure. First, the signal is deserialized to four half rate signals, which are then decoded using two half rate XOR gates. Afterwards, these two streams are further deserialized and provided to the four quarter rate output buffers. The techniques used in this chipset led to three patent applications on the chip implementation.

To perform system tests with the designed chipset, a test board that does not limit the chip's performance is required. To this end, several connector footprints and routing techniques were developed and compared. The results were presented at the IEEE Workshop on Signal and Power Integrity (SPI) 2014. Using 1.85 mm Rosenberger connectors, traces with smooth frequency response up to 67 GHz are shown, including tapering down to the chip inputs and transitioning from coupled to uncoupled transmission

lines. This technique, together with thermosonic flip chip mounting of the chips on the board, led to multiple successful test boards, ranging from a small 4 layer Rogers RO4003C board with only one high speed chip, to a large 24 cm x 30 cm 12 layer Megtron 6 board with two receiver and two transmitter chips.

Combining the designed receiver chip and the four layer test board, an 84 Gb/s transmission link is achieved across 20 cm of coax cable connected in between the receiver and transmitter test boards with a BER $< 10^{-11}$ without any form of forward error correction. The total channel loss at 42 GHz is approximately 14 dB with the channel equalized to the duobinary shape by the 5 tap 12.4 ps spaced feedforward equalizer available at the transmitter. Up to now, this is the fastest duobinary modulated serial communication link which subsequently led to a journal paper published in Electronic Letters.

Finally, a complete backplane test bed was built to show the chip's performance at 56 Gb/s. The test bed contains two active daughter cards, each with 4 chips, and a Megtron 6 ExaMAX backplane. Error free (BER $< 10^{-13}$) transmission was shown across backplane links up to 50 cm with more than 40 dB loss at 28 GHz. Transmission across a channel with 33 dB of loss at 28 GHz was demonstrated live at DesignCon 2015 in Santa Clara (USA), in the booth of FCI with a 28 Gb/s cross talk aggressor. A paper on this work was presented at DesignCon 2015 and won a best paper award.

The work presented in this dissertation gained a lot of attention from the industry during the Bell Labs Future X days, and during DesignCon 2015. The developed chipset and board design techniques can be used for many different applications, such as: driving short high-speed optical channels, transmission across complex high-loss backplane channels or short-range chip-to-module applications.

Together with three other researchers from the INTEC Design group, we aquired project funding from the Industrial Research Fund of Ghent University to further refine the technology developed during this work and prepare it for commercialization.

List of Publications

Publications in International Journals

T. De Keulenaer, G. Torfs, Y. Ban, R. Pierco, R. Vaernewyck, A. Vyncke, Z. Li, J.H. Sinsky, B. Kozicki, X. Yin, and J. Bauwelinck. 84 Gb/s SiGe BiCMOS duobinary serial data link including serialiser/deserialiser (serdes) and 5-tap FFE. *Electronics Letters*, 51(4):343–345, 2015

Y. Ban, T. De Keulenaer, G. Torfs, J.H. Sinsksy, B. Kozicki and J. Bauwelinck. Experimental evaluation of NRZ and duobinary up to 48 Gb/s for electrical backplanes. *Electronics Letters*, 51(8):617–619, 2015

R. Pierco, G. Torfs, T. De Keulenaer, B. Vandecasteele, J. Missinne, and J. Bauwelinck. A ka-band SiGe BiCMOS power amplifier with 24 dBm output power. *Microwave and Optical Technology Letters*, 57(3):718–722, 2015

K. De Kerpel, T. De Keulenaer, S. De Schampheleire, and M. De Paepe. Capacitance sensor measurements of upward and downward two-phase flow in vertical return bends. *International Journal of Multiphase Flow*, 64:1–10, 2014

Publications in International Conferences

T. De Keulenaer, J. De Geest, G. Torfs, J. Bauwelinck, Y. Ban, J.H. Sinsky, and B. Kozicki. 56+ Gb/s serial transmission using duobinary signaling. *Proc. IEC DesignCon (Santa Clara, CA, 2015)*, 2015

T. De Keulenaer, G. Torfs, R. Pierco, and J. Bauwelinck. A digitally controlled threshold adjustment circuit in a 0.13μm SiGe BiCMOS technology for receiving multilevel signals up to 80 Gb/s. In *2014 10th Conference on*

Ph. D. Research in Microelectronics and Electronics (PRIME), pages 1–4. IEEE, 2014

T. De Keulenaer, Y. Ban, Z. Li, and J. Bauwelinck. Design of a 80 Gb/s SiGe BiCMOS fully differential input buffer for serial electrical communication. In *2012 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 237–239. IEEE, 2012

T. De Keulenaer, Y. Ban, G. Torfs, S. Sercu, J. De Geest, and J. Bauwelinck. Measurements of millimeter wave test structures for high speed chip testing. In *2014 IEEE 18th Workshop on Signal and Power Integrity (SPI)*, pages 1–4. IEEE, 2014

R. Pierco, T. De Keulenaer, G. Torfs, and J. Bauwelinck. Analysis and design of a high power, high gain SiGe BiCMOS output stage for use in a millimeter-wave power amplifier. In *2014 10th Conference on Ph. D. Research in Microelectronics and Electronics (PRIME)*, pages 1–4. IEEE, 2014

A. Dierck, T. De Keulenaer, F. Declercq, and H. Rogier. A wearable active gps antenna for application in smart textiles. In *32nd ESA Antenna Workshop on Antennas for Space Applications: From technologies to architectures*, 2010

K. De Kerpel, T. De Keulenaer, S. De Schampheleire, and M. De Paepe. Two-phase flow behaviour in a smooth hairpin tube: analysis of the disturbance using capacitive measurements. In *10th International Conference on Heat Transfer, Fluid Mechanics and Thermodynamics*, pages 1314–1321, 2014

K. De Kerpel, C. De Groof, M. Dreesen, T. De Keulenaer, and M. De Paepe. Two phase flow behaviour in a smooth hairpin tube. In *4th IIR Conference on Thermophysical Properties and Transfer Processes of Refrigerants*, pages 227–234, 2013

K. De Kerpel, C. De Groof, M. Dreesen, T. De Keulenaer, and M. De Paepe. Capacitive sensor measurements on two-phase flow in smooth return bends. In *8th world conference on experimental heat transfer, fluid mechanics and thermodynamics*, 2013

Publications in National Conferences

T. De Keulenaer, J. Vandewege, and J. Bauwelinck. 40 Gb/s transceiver design. In *12th FEA PhD Symposium, Interactive poster session*, Ghent, Belgium, December 2011

Patents

J.H. Sinsky, G. De Peuter, G. Torfs, Z. Li, and T. De Keulenaer. Circuitry and method for multi-level signals. European patent application, EP14305284.3, filed in March 2014

J. Bauwelinck, G. Torfs, Y. Ban, and T. De Keulenaer. Improvements in test circuits. European patent application, EP14161772.0, filed in March 2014

T. De Keulenaer, R. Vaernewyck, J. Bauwelinck, and G. Torfs. Improvements in or relating to signal processing. European patent application, EP14161804.1, filed in March 2014

1 Introduction

THIS dissertation pertains to high-speed backplane communication technologies. To introduce the basic concepts, this chapter discusses where a backplane can be found in today's telecom and datacom systems and how backplanes are evolving to reach next generation speed requirements. This is done by giving an overview of the different backplane topologies and architectures as well as an overview of the different modulation formats used.
After providing this necessary background, an overview of the work and an outline of the dissertation is presented.

1.1 Background

Starting in the late nineties the Internet became increasingly present in people's daily lives. From 2000 onwards more and more people had home access to high-speed broadband internet. In the mean time, they grew so accustomed to it that it became indispensable. Between 2000 and 2010 the broadband speeds grew from around 1 Mb/s to above 100 Mb/s. Only a few years later, mobile internet usage became more and more popular and providers upgraded their systems from 3G to 4G systems, allowing mobile speeds of above 50 Mb/s needed for video streaming, gaming, communication, etc. This impressive evolution in access speeds leads to new bottlenecks deeper in the network. All data coming from millions of users has to

be processed and transferred by the service providers or in the data center. The advances in the data rate of modern communication systems has led to the need for an increased speed in inter-chip communication over so-called backplanes.

These inter-chip serial communication links are becoming more ubiquitous in electronic system designs at an astonishing rate. Today's typical backplane systems use multiple daughter cards communicating over a backplane using non return to zero on-off keying (NRZ OOK) with serial links at 25 Gb/s or below. To move forward and use higher serial speeds, system designers will have to deviate from the traditional ways and think about more advanced modulation formats and/or different system architectures [1].

1.1.1 Backplane configurations

A typical backplane setup, of which a photo is shown in Figure 1.1, consists of two or more daughter cards interconnected via a backplane. Due to the excessive amount of frequency dependent loss found in these kind of structures, alternatives are approaching the market for high-speed backplanes (28+ Gb/s). Such setups are, among others, cable backplanes, midplane systems and direct orthogonal systems. Another alternative is the use of shorter backplane channels to reduce the loss at higher rates.

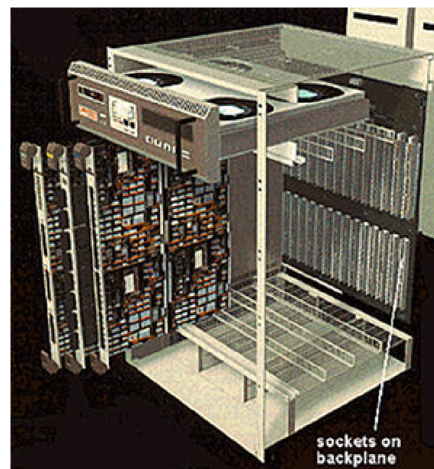


Figure 1.1: Photo of a typical backplane system, a Cabletron MMAC PLUS network hub.

First an overview of the possible backplane topologies, shown in Figure 1.2, will be given.

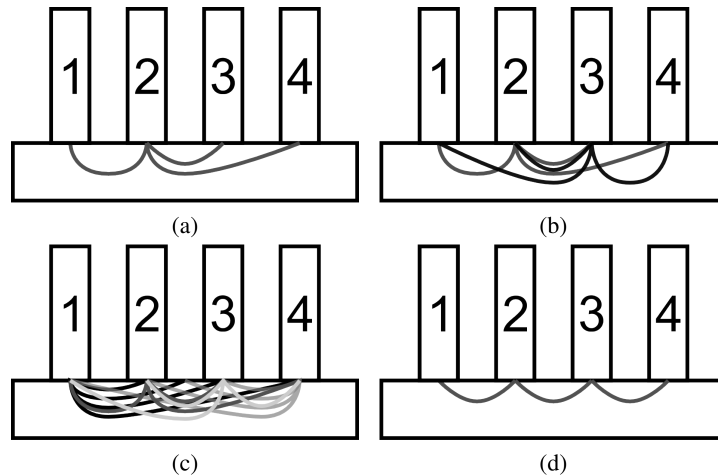


Figure 1.2: Schematical backplane topologies: Single Star topology (a), Dual Star topology (b), Full Mesh topology (c) and a Chain topology (d).

Single Star topology

In network setups today, the star architecture is one of the most common high-speed serial topologies. The advantage is that it reduces the chance of network failure by connecting all the systems to a central node. A failure of a link from any peripheral node to the central node results in the isolation of that peripheral node from all others. As a result, the rest of the system remains unaffected [2].

A basic form of a star topology is shown in Figure 1.2a. In this implementation the central node (node 2) is typically aggregating all traffic from the peripheral nodes (nodes 1, 3 and 4).

The disadvantages of a single star topology are the high dependency on the central node and the large line lengths due to the need to connect to each peripheral node. In case of failure of the central node the system becomes useless.

Dual Star/ Multi-Star topology

To reduce the dependency on the central node in a single star topology, two or more central nodes can be used, as illustrated in Figure 1.2d. Introducing multiple central nodes provides redundancy in

mission critical system applications in case of failure, or allows to upgrade the node hardware.

Mesh topology

A fully connected mesh topology, when applied to a backplane application, does not have one or more central nodes as in the case of star topologies, as illustrated in Figure 1.2b. Instead, each node connects with all other nodes forming a mesh. Its major disadvantage is the number of connections, which grows significantly with the number of nodes. This requires additional backplane connector pins and layers to interconnect them. Because of this, it is impractical for large systems and only used when there are a small number of cards that need to be interconnected.

Chain topology

When moving to higher speeds, line length is sometimes more critical than redundancy. Therefore a chain topology as illustrated in Figure 1.2d can be of interest. Each node is only connected to its neighbors allowing minimal line lengths.

The main drawback of this topology is that the failure of one node leads to the failure of a large part of the system. To reduce this dependency one can utilize a chain topology in which each node is connected to its neighbors and their neighbors. This kind of enhanced chain topology increases the reliability and still allows for a reduction in line length.

Now an overview of the possible backplane architectures shown in Figure 1.3 will be given.

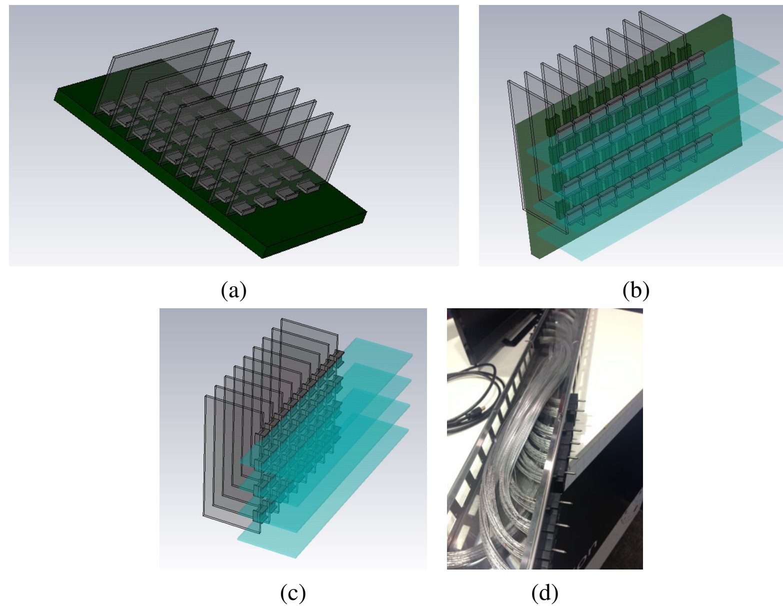


Figure 1.3: Schematical backplane architectures: (a) typical backplane setup, (b) orthogonal setup using midplane, (c) a direct mate orthogonal setup and (d) a cable backplane setup.

Traditional backplane

In traditional backplane implementations the daughter cards are oriented parallel to each other and mounted on a backplane, as shown in Figure 1.3a.

Orthogonal/midplane/DMO backplane

One way to reduce the frequency dependent loss is to shorten the link or reduce the amount of connectors and connector footprints. This can typically be seen in orthogonal setups where daughter cards are oriented perpendicular to each other, as shown in Figure 1.3b. In most cases there is a midplane in between, however, one can increase the performance even more by removing this midplane and using direct mate orthogonal connectors (DMO), illustrated in Figure 1.3c.

Cable backplane

Another way to reduce the frequency dependent loss without decreasing the physical distance between the daughter cards is using cables

instead of a backplane. A photo of a cable backplane is shown in Figure 1.3d. Cables typically consist of a dielectric that has better high-frequency performance. Another benefit of using cables is the routing: since all used cables are shielded twinax cables, one can route in as many layers as needed. Therefore, the routing can be changed after installation without having to replace the complete backplane board.

1.1.2 Modulation formats

In this subsection the main baseband modulation formats will be discussed: NRZ, 4 level Pulse Amplitude Modulation (PAM4) and duobinary (DB). These are the most realistic formats for high-speed low-cost short-reach wireline and optical systems. Later in this dissertation some other formats will be mentioned and explained. Each of these three formats has its drawbacks and benefits. In Figure 1.4 the three formats are shown next to each other.

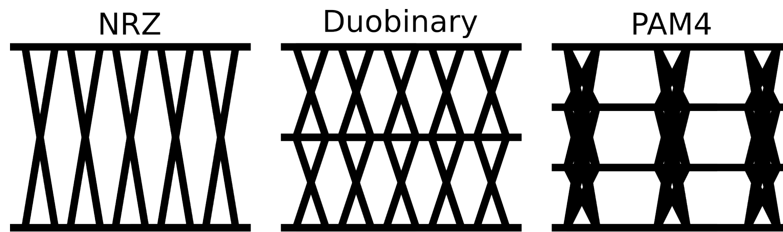


Figure 1.4: Overview of backplane modulation formats, from left to right: NRZ, duobinary and PAM4.

NRZ

NRZ excels in its low complexity and, therefore, native power consumption. However, due to the excessive bandwidth requirement, it needs the strongest equalization of all modulation formats under comparison. NRZ typically needs Forward Error Correction (FEC) to provide reliable communication, which increases power consumption and latency.

PAM4

PAM4 combines two bits in each symbol, reducing the symbol rate by a factor of two. This goes hand in hand with the fact that the bandwidth is halved. However, it also adds complexity, increases receiver sensitivity requirements and needs a linear output at the transmitter. Additionally, it adds latency in the decoding and FEC difficulties.

Duobinary

Duobinary uses known inter-symbol interference to reduce the required channel bandwidth, which makes it a specific form of partial response signaling. This means that two consecutive bits are merged together in a known way. To prevent error propagation, a precoder is added. Merging two consecutive bits together can be done using a feed forward equalizer (FFE). The same equalizer that would be used for NRZ could be used using different parameters to convert the combined response of the FFE and the channel to a duobinary channel instead of a channel without inter-symbol interference (ISI) (as would be the case for NRZ or PAM4 optimization). This means that duobinary can actually embrace part of the loss of the channel and use it to obtain its typical 2-eye duobinary shape. At the receiver side, the sensitivity needs to be higher compared to that of an NRZ receiver and the input needs to be linear. However, the sensitivity requirements are less stringent with respect to a PAM4 receiver since there are two eyes instead of three with a PAM4 signal. The decoding of the duobinary stream to an NRZ stream can be done using a rectifier, which in modern designs is typically implemented using the semi-digital approach proposed by Sinsky [3] consisting of an XOR gate combined with two level shifters. An additional advantage of duobinary is that do to the nature of correlative coding, some error detection/correction can be done as explained in [4]. This is however not commonly integrated yet in high speed duobinary receivers.

Going into more detail, comparing the bandwidth requirements, we can check the cumulative power spectral density of these modulation schemes. From Figure 1.5 it is clear that PAM4 and duobinary need considerably less bandwidth compared to NRZ. Selecting 90% bit energy (of a square pulsed modulated signal) as the minimum bandwidth necessary results in 35% of the bitrate for PAM4 and duobinary and 66% of the bitrate for NRZ, which clearly indicates the bandwidth reduction obtained by using PAM4 or DB versus NRZ.
Another more common way for system architects to compare different modulation formats is the Nyquist frequency (fn). Fn is defined as being the highest frequency minimally required to receive the transmitted stream without unwanted ISI. An overview for each modulation format is shown in Figure 1.6. For NRZ and PAM4 these are well known and equal to half and a quarter of the bitrate respectively. For duobinary it is less intuitive but as depicted in Figure 1.6, the Nyquist frequency is at 1/3 of the bitrate. To compare the modulation formats, each format has a penalty with

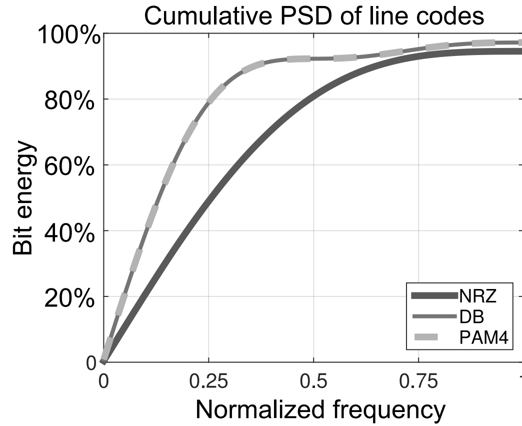


Figure 1.5: Cumulative power spectral density (PSD) plot comparing NRZ, duobinary modulation and PAM4. The bit energy of PAM4 and duobinary modulation is concentrated in a lower part of the spectrum compared to that of NRZ.

respect to the achieved eye height. This way one can determine which modulation format would have the largest vertical eye opening for a given channel. For PAM4 this signal to noise penalty is well defined at 9.54 dB ($-20log_{10}(1/3)$) since the eye heigth is one third of the amplitude at frequency fn_{PAM4}. For duobinary the eye height is three quarters of the amplitude at fn_{DB} resulting in a signal to noise penalty of only 2.5 dB ($-20log_{10}(3/4)$). Although fn_{DB} is 33% higher compared to fn_{PAM4} there is an SNR penalty difference of 7 dB. Which is comparable to the measured 6 dB difference in [5].
In summary, three situations can be distinguished when comparing channels with a linear loss profile (e.g. a trace on a printed circuit board) and only taking the eye height into account, a typical example is given in Figure 1.7. For this kind of channel that the slope equals IL_{PAM4}/fn_{PAM4} as well as IL_{DB}/fn_{DB} and IL_{NRZ}/fn_{NRZ}. Taking into account the SNR penalty one can calculte that as long as the channel loss is less than 3.75 dB at the quarter bit rate (fn_{PAM4}) one should use NRZ for maximal eye opening. Once the loss at fn_{PAM4} is between 3.75 dB and 21 dB, duobinary is the preferred modulation format and beyond 21 dB loss at fn_{PAM4} PAM4 will achieve the highest vertical eye opening.
It should be noted that in current systems an insertion loss of 21 dB at fn_{PAM4} (or for linear channels 42 dB at fn_{NRZ}) is very high and only occurs when working with cheap material or at very high bitrates. Also many of the current channels are not linear up to the fn_{NRZ}. Further in

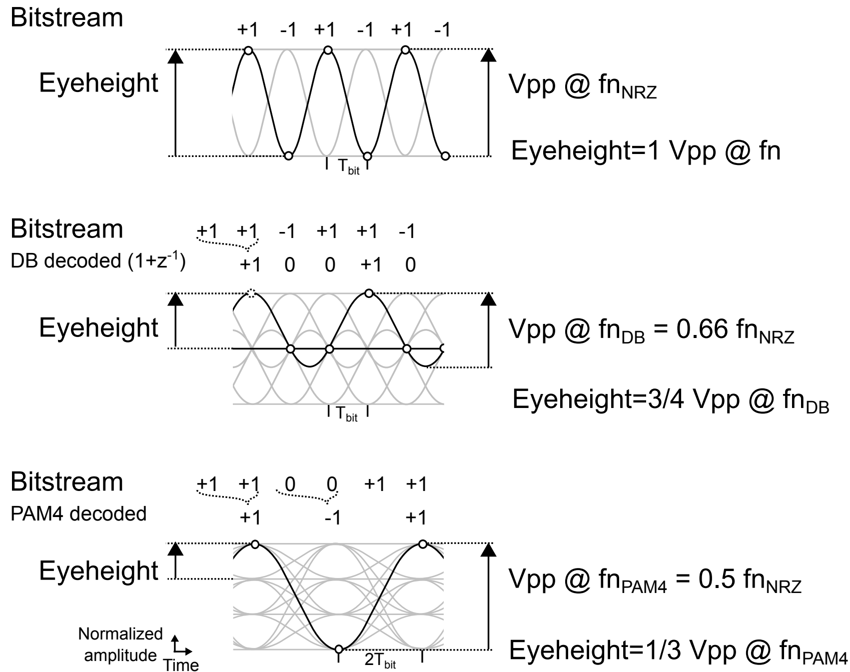


Figure 1.6: Time domain represenation of the Nyquist frequency and the associated eyeheight versus peak ampltude ratio of NRZ (top), Duobinary modulation (middle) and PAM4 (bottom).

this dissertation, the loss of a channel will be specified at the fn_{NRZ} for easy comparison to other commercial channels (unless noted otherwise).

Another thing that is clarified in Figure 1.6 is the horizontal eye opening, when using minimum bandwidth pulses. It is clear that, although PAM4 has the lowest Nyquist frequency and needs only half the sample rate of NRZ and duobinary, the horizontal eye opening of a PAM4 signal is significantly reduced when reducing the channel bandwidth to fn_{PAM4}. This results in a comparable jitter tolerance with respect to NRZ and DB, as opposed to what one would expect from halving the baudrate.

1.1.3 Duobinary system

In Figure 1.8 a typical example of a duobinary system is shown, assuming 84 Gb/s duobinary transmission. Duobinary is a special case of partial response NRZ coding [6] as explained earlier. One of the main advantages is the relatively easy receiver chain, however, the drawback is that you need

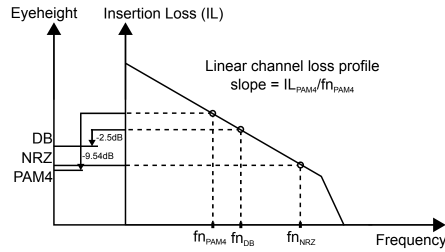


Figure 1.7: Example of determining the largest vertical eye opening for NRZ, duobinary modulation and PAM4 on a channel with a linear loss profile.

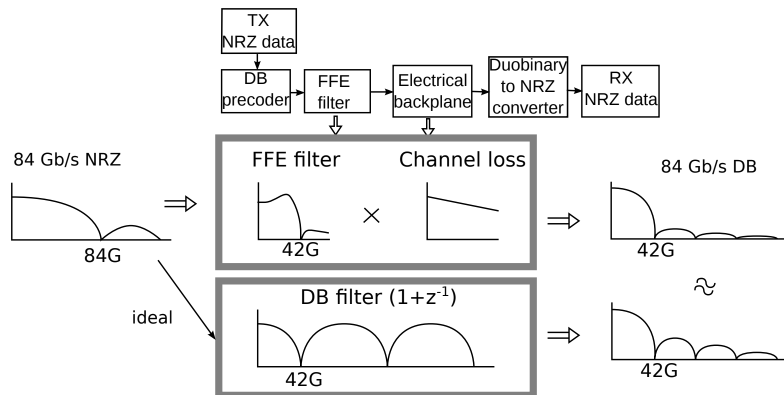


Figure 1.8: Overview of the duobinary signal chain.

a precoder to prevent error propagation. At the transmitter side, a data stream is precoded and shaped to produce the wanted duobinary shape at the output of the channel. Hence, the transmitter contains a precoder and a shaping filter.

The precoding can be implemented in different ways and can also be done on a sub rate if one decides to use multiple sub rate data streams at the input of the transmitter. On the full rate it can be implemented as a toggle flipflop, which toggles the signal when the input is high and keeps the output when the input signal is low, or using an XOR gate as depicted in Figure 1.9a. On a sub rate a duobinary precoder can be implemented as described in [7] and illustrated in Figure 1.9b.

The shaping filter can also be implemented in different ways e.g. a Continues Time Linear Equalization filter (CTLE) or a Feed Forward Equalizer (FFE). It is intended to shape the signal in such a way that in combination with the channel, a low pass filter is obtained, with transfer function, $H(z) = 1 + z^{-1}$. This results in a three level signal at the receiver input.

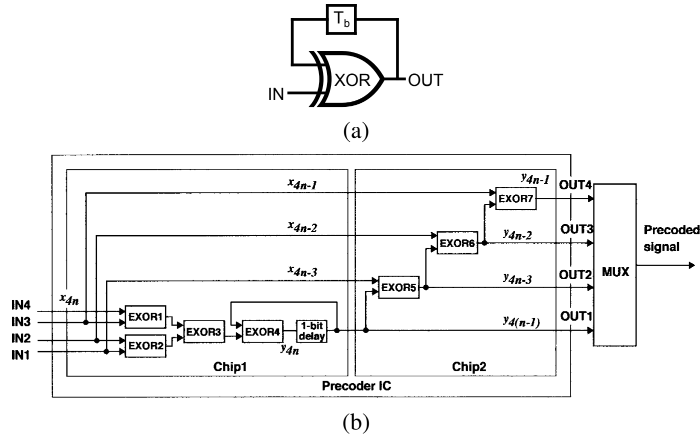


Figure 1.9: Example of a duobinary precoder implementation: a typical full rate precoder (a) and a quarter rate precoder (b).

At the receiver side there is a duobinary-to-binary converter, which can be implemented as a rectifier or its semi-digital equivalent, a XOR gate as shown in Figure 1.10 and explained in more detail in [3, 8]. More advanced receiver structures are also possible, making optimal use of duobinary and its error detecting properties. A more in depth overview of how a duobinary receiver can be implemented on-chip can be found in Chapter 2. To

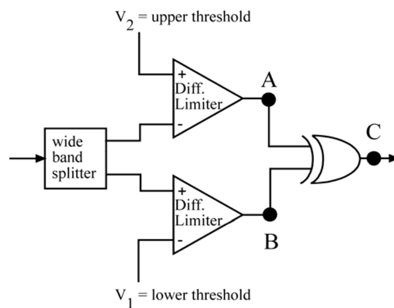


Figure 1.10: Decode duobinary using an XOR gate. (from [8])

implement the use of duobinary signaling in a real-life system the use of a CDR is inevitable. In litrature some topologies are explained implemening a CDR to receive duobinary modulated signals either by using NRZ based techniques after the decoding or by using multilevel techniques [9–11].

1.2 Overview of the work

This dissertation is based on the design and testing of a high-speed duobinary receiver I have worked on over the last 5 years at the Design laboratory of the department of information technology (INTEC) at Ghent University.

This work in particular was based on my research performed during his personal IWT grant and in close collaboration with the IWT O&O project ShortTrack together with Alcatel Lucent Bell Labs[1]. His personal IWT grant focusses on design methodologies and new chip architectures to scale to rates above 40 Gb/s for relatively long backplane links (order 0.5 m). During the ShortTrack project an implementation of this architecture was done with the goal of reaching 100 Gb/s across 10 cm backplane links. After the ShortTrack project I was also involved in a follow up collaboration with FCI[2], a to provide a live demonstrator at their booth during DesignCon 2015 showing error free 56 Gb/s transmission over Megtron 6 backplane links.

My contributions during these projects consisted of the technology explorations, the flip-chip assembly and the optimization methodology for the high-speed transceiver chips, followed by the design of the 100 Gb/s duobinary front end and the design of the high speed test boards for system tests and demonstrations.

1.2.1 Chip design and analysis

During this PhD research, a duobinary receiver was designed. Two chip design cycles were needed to realize a fully functional prototype. The first run was used to test different building blocks such as the input buffer and the duobinary front end. The second tape-out contained an optimized duobinary front end based on the results of the first chip. Also a back end which recovers the original NRZ data out of the two duobinary eyes was added. The back end, consisting of an XOR gate implemented in a deserializer was designed together with our colleague Zisheng Li.

All building blocks were first designed and optimized separately and afterwards combined to larger sub structures. These structures such as a level shifting limiting amplifier consisting of different gain and level shifting stages were then again optimized together to obtain optimal performance.

[1]Alcatel Lucent Bell Labs: Inspiring innovation, changing the way you see the world. http://www.alcatel-lucent.com/bell-labs

[2]FCI: Setting the Standard for Connectors. http://www.fci.com

Due to the high bitrates and accompanying bandwidths it became clear in the first years of this work that the board design and the connector and mounting methods would have a significant impact on the system performance. One could even say that the test board design becomes equally important as the chip design.

1.2.2 Measurements and results

After the designed duobinary receiver was fabricated and mounted on different test boards, multiple experiments were conducted. The main conclusions and results of these experiments are bundled in chapters 4 and 5 based on publications [12, 13]. Most of the experiments were done in close collaboration with Yu Ban and later also with Joris Van Kerrebrouck.

1.3 Outline of the dissertation

The dissertation consists of an introduction in Chapter 1 followed by 2 parts. Part I contains Chapter 2 and Chapter 3 and treats the design and analysis. Part II consists of Chapter 4 and Chapter 5 and focusses on measurement results. The last chapter formulates a conclusion on this topic after almost five years of work. An overview of the content of the succeeding chapters is given below.

Chapter 2
The second chapter focusses on the chip design performed during this thesis. In the first part an overview is given of the implemented high-speed duobinary receiver, indicating all high-speed building blocks of the duobinary front end and the DEMUX/XOR. The second part [14] focusses on the input buffer of the receiver chips and how it provides the necessary matching, gain and linearity. It explains how the use of a transimpedance amplifier matched to a $100\,\Omega$ differential impedance can be a very effective wideband low noise input buffer. Subsequently, the level shifter is explained [15]. It shows how the duobinary signal is shifted so that either the upper or the lower eye can be received. To obtain a high shifting resolution a digital current control is implemented. The additional high-speed building blocks are explained afterwards, giving the reader an overview of all main components included in the receiver chip. The chapter concludes with an overview of the power consumption of the chip and an overview of the achieved results.

Chapter 3
This chapter (see [16]) deals with the characterization of different millimeter wave structures, and the design of high-speed test boards. It focusses on the importance of connector footprint design as well as trace and via impedance matching. Different connectors are compared and multilayer board buildups are explained. An overview is presented of different commercially available flip chip mounting techniques to allow for broadband interconnection bandwidths above 50 GHz. These were necessary for the successful demonstration of the designed chips.

Chapter 4
The focus in this chapter is on a 84 Gb/s serial link measurement that was performed [12]. It explains the way the transmitter is able to shape a high-bandwidth, lossy channel to a duobinary shape and how the experiment was set up to reach this record breaking rate in serial electrical transmission. Due to the limited availability of the necessary equipment we were only able to set up the 84 Gb/s link across a coaxial link and were time limited to try other interesting channels.

Chapter 5
In the penultimate chapter (see [13]) the focus lies on backplane communication links. The first part yields an overview of the current and future standards of backplane communication. This is followed by an introduction to duobinary signaling and an overview of the custom ASIC design. The second part discusses frequency as well as time domain measurements on the test setup and shows how the boards designed in Chapter 3 and the chip designed in Chapter 2 are used. The measurements shown in this chapter were also repeated during a live demo at DesignCon 2015 in the booth of FCI.

Chapter 6
The final chapter summarizes the progress made in high-speed serial communication links and highlights the most important results. Some suggestions are listed for further research and valorization of the results and know-how obtained during this PhD.

References

[1] W. Beyene, YC. Hahm, D. Secker, and D. Mullen. Design and characterization of interconnects for serial links operating at 56 Gb/s and beyond. *Proc. IEC DesignCon (Santa Clara, CA, 2015)*, 2015.

[2] B. Simonovich. Backplane architecture and design. 2011.

[3] J.H. Sinsky, M. Duelk, and A. Adamiecki. High-speed electrical backplane transmission using duobinary signaling. *IEEE Transactions on Microwave Theory and Techniques*, 53(1):152–160, 2005.

[4] A. Lender. Correlative digital communication techniques. *IEEE Transactions on Communication Technology*, 12(4):128–135, 1964.

[5] Yamaguchi, K.; Sunaga, K.; Kaeriyama, S.; Nedachi, T.; Takamiya, M.; Nose, K.; Nakagawa, Y.; Sugawara, M.; Fukaishi, M. J. Lee, MS. Chen, and HD. Wang. 12 Gb/s duobinary signaling with x2 oversampled edge equalization. In *2005 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, IEEE, 2005.

[6] M. Hossain and AC. Carusone. Multi-Gb/s bit-by-bit receiver architectures for 1-d partial-response channels. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(1):270–279, 2010.

[7] M. Yoneyama, K. Yonenaga, Y. Kisaka, and Y. Miyamoto. Differential precoder ic modules for 20-and 40-Gb/s optical duobinary transmission systems. *IEEE Transactions on Microwave Theory and Techniques*, 47(12):2263–2270, 1999.

[8] J. Lee, MS. Chen, and HD. Wang. Design and comparison of three 20-Gb/s backplane transceivers for duobinary, pam4, and nrz data. *IEEE Journal of Solid-State Circuits*, 43(9):2120–2133, 2008.

[9] K. Sunaga, H. Sugita, K. Yamaguchi, and K. Suzuki. An 18Gb/s duobinary receiver with a CDR-assisted DFE. In *2009 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, IEEE, 2009.

[10] L.P. Baskaran, A. Morales, and S. Agili. Transmitter Pre-emphasis and Adaptive Receiver Equalization for Duobinary Signaling in Backplane Channels. In *Digest of Technical Papers. International Conference on Consumer Electronics (ICCE)*, 2007.

[11] M. Hsieh, and G.E. Sobelman. SiGe BiCMOS PAM-4 clock and data recovery circuit for high-speed serial communications. In *Digest of Technical Papers. IEEE International Systems-on-Chip Conference*, 2003.

[12] T. De Keulenaer, G. Torfs, Y. Ban, R. Pierco, R. Vaernewyck, A. Vyncke, Z Li, J.H. Sinsky, B. Kozicki, X. Yin, and J. Bauwelinck. 84 Gb/s SiGe bicmos duobinary serial data link including serialiser/deserialiser (serdes) and 5-tap ffe. *Electronics Letters*, 51(4):343–345, 2015.

[13] T. De Keulenaer, J. De Geest, G. Torfs, J. Bauwelinck, Y. Ban, J. Sinsky, and B. Kozicki. 56+ Gb/s serial transmission using duobinary signaling. *Proc. IEC DesignCon (Santa Clara, CA, 2015)*, 2015.

[14] T. De Keulenaer, Y. Ban, Z. Li, and J. Bauwelinck. Design of a 80 Gb/s SiGe bicmos fully differential input buffer for serial electrical communication. In *2012 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 237–239. IEEE, 2012.

[15] T. De Keulenaer, G. Torfs, R. Pierco, and J. Bauwelinck. A digitally controlled threshold adjustment circuit in a 0.13μm SiGe bicmos technology for receiving multilevel signals up to 80 Gb/s. In *2014 10th Conference on Ph. D. Research in Microelectronics and Electronics (PRIME)*, pages 1–4. IEEE, 2014.

[16] T. De Keulenaer, Y. Ban, G. Torfs, S. Sercu, J. De Geest, and J. Bauwelinck. Measurements of millimeter wave test structures for high speed chip testing. In *2014 IEEE 18th Workshop on Signal and Power Integrity (SPI)*, pages 1–4. IEEE, 2014.

Part I

Design and analysis

2
Receiver design

Timothy De Keulenaer, Ramses Pierco, Yu Ban, Zhisheng Li, Guy Torfs and Johan Bauwelinck

Based on the publications presented by Timothy De Keulenaer at ICECS 2012 and PRIME 2014 extended with an introduction and a description of the remaining subblocks.

Section 2.2 was presented at the 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2012)
Section 2.3 was presented at the 10th Conference on Ph.D Research in Microelectronics and Electronics (PRIME 2014)

⋆ ⋆ ⋆

THE authors would like to thank the Agency for Innovation by Science and Technology in Flanders (IWT) and Integrand Software for access to their EM solver EMX. As well as Ricardo Andres Aroca and Shahriar Shahramian from Alcatel Lucent for the thorough schematic and layout reviews.

2.1 Introduction and design overview

In this chapter the chip design activities are discussed. The chip designed in this work was created in close collaboration with the design of a matching transmitter chip described in the dissertation of my colleague Yu Ban. In this section the receiver topology will be explained together with the importance of all subblocks. The main blocks were published on international conferences and a reformatted version of these publications is imported in sections 2.2 and 2.3. Other blocks will briefly be described in the succeeding sections.

Figure 2.1 shows the receiver topology with at the left a low-noise amplifier (LNA) implemented as a transimpedance amplifier (TIA). The TIA is the first active stage on the chip directly connected to the chip's bump pads and is followed by a set of level shifters, to extract either the upper or the lower eye of the duobinary modulated signal. The level shifters are contained in the level shifting limiting amplifier (LSLA) which amplifies and shifts the upper or the lower eye until a logic-compatible digital swing is obtained. The LSLA outputs are connected to two signal paths: the front end test port and the DEMUX/XOR circuit with the possibility to turn off the test ports via a serial peripheral interface (SPI).

Implementing duobinary signalling at high speeds using the traditional wide band rectifier is not possible using today's technologies. The method proposed by Sinsky in [1] and shown in Chapter 1 is a lot more appropriate for a high-speed implementation. However, each stage adds jitter. At the input of the receiver, the jitter on the duobinary modulated eye is typically 5 to 6 ps. A period of 10 ps for a 100 Gb/s signal severely limits the remaining jitter budget.

To reduce the added jitter, this architecture resamples the signal as soon as possible. This resampling operation reduces the jitter to the clock jitter but can introduce errors if the phase of the clock is misaligned with respect to the data.
The main block in the proposed receiver architecture that could benefit from a clocked input is the XOR gate. Since full rate sampling beyond 80 Gb/s was not possible in the selected technology, the signal after the LSLA is first down-sampled to a half rate signal and a XOR operation is performed afterwards, as shown in Figure 2.1. This architecture is also the subject of a patent application that has been filed (EP14305284.3).

Implementing the duobinary receiver in this way diminishes the jitter intro-

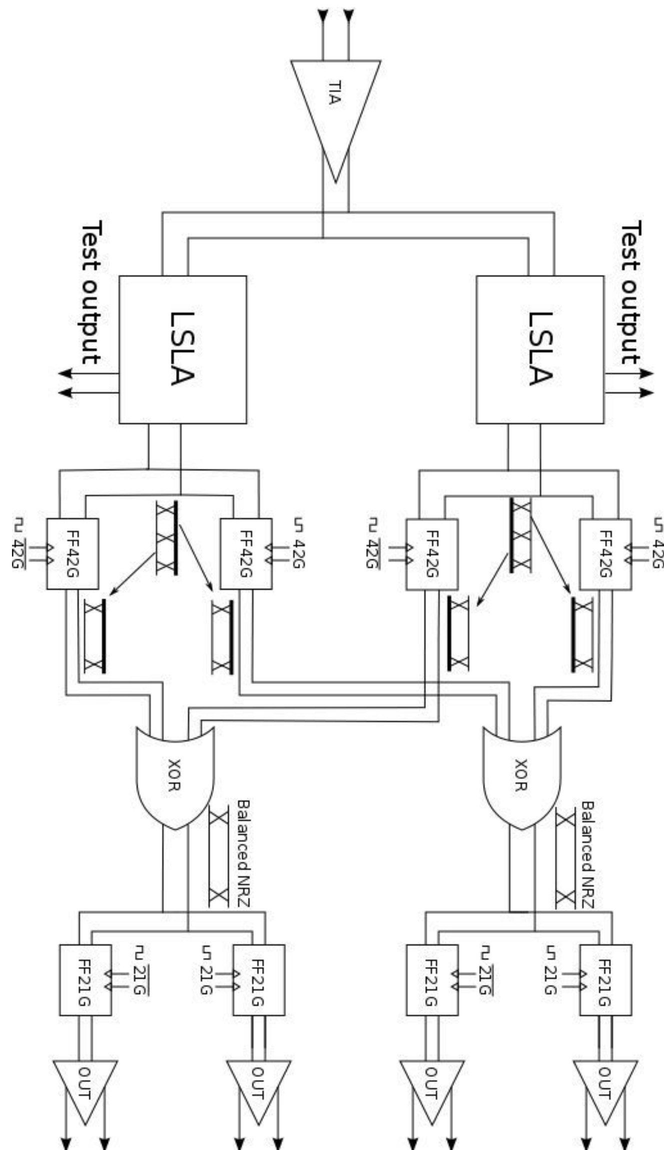


Figure 2.1: Implemented receiver architecture for receiving a 84 Gb/s duobinary signal.

duced by the XOR and runs the XOR at half the channel bit rate. However, a synchronized clock at the receiver side is necessary. This clock can be generated and phase shifted on chip using a CDR circuit, however this is outside of the scope of this dissertation.

2.2 Design of a 80 Gbit/s SiGe BiCMOS fully differential input buffer for serial electrical communication

2.2.1 Introduction

At high data rates, inter-chip electrical communication over PCB traces or long cables becomes challenging due to excessive frequency dependent channel attenuation, which causes large amounts of inter-symbol interference (ISI). Combined with impedance mismatch, this results in a very challenging environment for high-speed and high-bandwidth communication. The input stage of the receiver is optimized for a low noise figure (NF) and good input matching, but without sacrificing area, gain and bandwidth.
In this paper, a linear input buffer designed in a 130 nm SiGe BiCMOS process is discussed. The designed buffer has >35 GHz of bandwidth and can be used for a multitude of modulation formats, e.g. NRZ, duobinary modulation, PAM4. Section 2.2.2 introduces the serial communication link and shows the focus of this paper. The topology used to design the matched input buffer is shown in Section 2.2.3, followed by a section on the electromagnetic simulations of long chip interconnections and the peaking inductor using EMX as an EM solver, which is based on the boundary element method [2]. Section 2.2.5 presents the results of the post layout simulations of the wideband input buffer for both AC and transient response. Section 2.2.6 compares the results achieved in this work with similar publications.

2.2.2 Block diagram

In Figure 2.2 an overview is given of a typical simplified serial communication link.

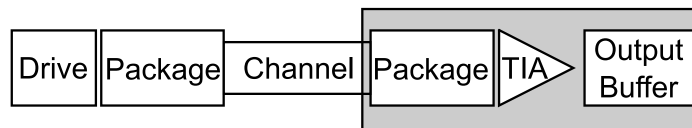


Figure 2.2: A schematic overview of a simplified link, the grey part is discussed in this paper.

The grey part of this block diagram will be discussed in this paper with emphasis on the TIA as a low noise input buffer for next generation serial electrical communication links.

2.2.3 Circuit description

2.2.3.1 Input buffer

A wide range of high-speed bipolar buffer topologies can be found in literature ranging from a simple differential pair or an emitter follower to more elaborate circuits optimized for noise matching or linearity using resistive degeneration [3]. In this work a transimpedance stage is implemented as shown in Figure 2.3. Recently, the use of shunt feedback to lower the optimal noise impedance to $50\,\Omega$ has been considered in low-noise voltage preamplifiers [4]. To obtain optimal broadband impedance matching to the system impedance Z_0, the open loop gain A and feedback resistor R_{fb} are chosen by:

$$Z_0 = R_{in} = \frac{R_{fb}}{1+A}. \tag{2.1}$$

Inductors are used to peak the open loop gain at the higher frequencies without excessively increasing the power dissipation. The design of these custom peaking inductors will be described in Section 2.2.4. When the loop gain reduces at higher frequencies, the input impedance changes. To reduce this effect a capacitor is added between the input and load impedance. The capacitor effectively cuts the loop and connects the input to the $50\,\Omega$ load impedance, this allows to increase the input matching bandwidth beyond the gain bandwidth. The feedback resistor is connected through emitter followers to the input to reduce the capacitive load on the collector resistors.

2.2.3.2 Output buffer

To characterize the TIA, a $50\,\Omega$ output buffer was designed using a typical current mode logic (CML) structure. At the input of this CML stage emitter followers were added to increase the input impedance and reduce the input common mode voltage. The schematic is shown in Figure 2.4.

2.2.4 EM modeling

Due to the high bandwidth (>35 GHz), little parasitics can be tolerated at the input. Wirebonding typically introduces 1 nH for each mm of wire and even with short bondwires it is hard to go below 0.3 nH. To solve this, flip chip bumps can be used. Furthermore, flip chip pads are typically smaller, compared to pand intended for wire bonding, which reduces the parasitic capacitance of the pad on the input. In 2002, a thorough study has been done by A. Jentzsch [5] which resulted in an overview table containing

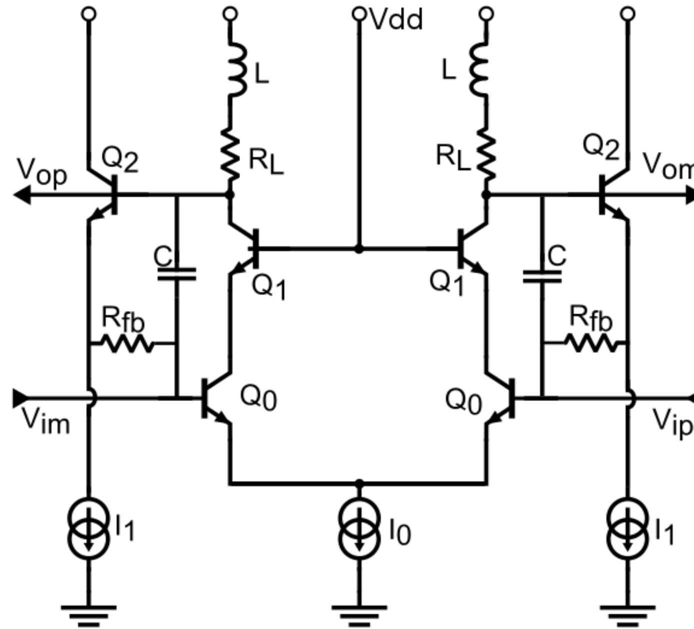


Figure 2.3: Fully differential input buffer. The load impedance R_L is $50\,\Omega$, the feedback impedance R_{fb} is $175\,\Omega$, the transistors are biased around maximum f_T.

different pad sizes and bump heights. Both measurements and EM simulation were done and compared in [5]. These results were verified using CST Microwave Studio[1], a 3D EM simulator, which shows that bandwidths exceeding 50 GHz are possible employing flip chip bumps.
Also due to the high bandwidth, accurate broadband models of long on-chip traces and on-chip inductors are required. These inductors were first designed as a separate component and later simulated as part of the TIA to take into account the influence of all parasitic elements. Simulations were performed using the EMX software of Integrand Software. As can be seen in the final layout (Figure 2.7) the size of the custom designed inductors is small compared to the bump pads.

2.2.5 Simulation results

The circuit of the input buffer was simulated including extracted layout parasitics and EM-models for the coils and long traces as well as equivalent models for the flip chip parasitics and ESD protection. The integrated

[1]CST MWS is a specialist tool for the 3D EM simulation of high frequency. http://www.cst.com

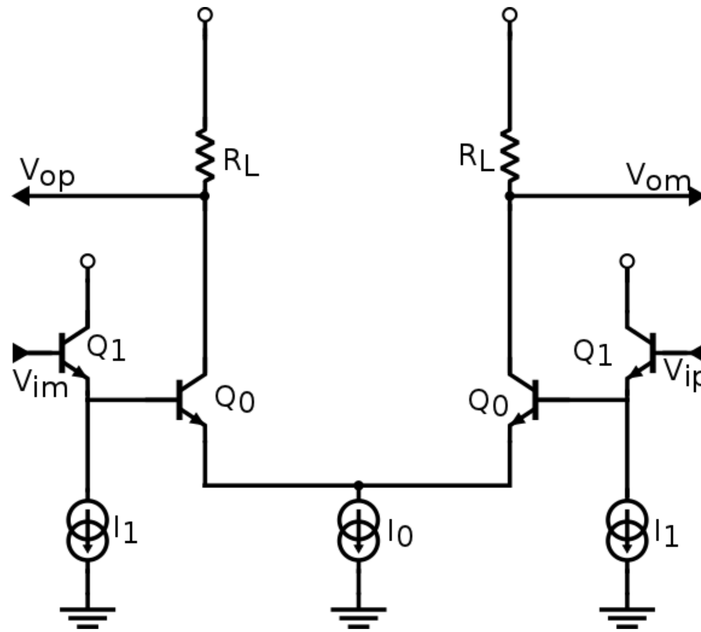


Figure 2.4: Schematic of the CML output buffer with emitter followers in front. Load impedance R_L *is* $50\,\Omega$*, current* I_0 *is* $6\,\mathrm{mA}$.

circuit is directly flip chipped on a circuit board.

2.2.5.1 Small-signal simulations

The S-parameters are shown in Figure 2.5. The amplifier has a gain of 14 dB and a 3-dB bandwidth of 37 GHz. The gain-bandwidth product is simulated to be 185 GHz. The simulated input return loss is below 14 dB within the 3-dB bandwidth and the output return loss is below 12.5 dB over the entire 3-dB bandwidth. The maximum noise figure over the 3-dB bandwidth is 8 dB and group delay variation is less than 3 ps.

2.2.5.2 Time-domain simulations

A transient simulation of the total circuit including flip chip parasitics, inductors and extracted capacitances is shown in Figure 2.6 with a 60 mVpp input signal at 80 Gb/s (12 ps bit period for NRZ and duobinary, 24 ps for PAM4) at 75 °C. The resulting NRZ eye shows a vertical eye opening of 200 mV and a horizontal eye opening of more than 10 ps with less than 2 ps of jitter. The duobinary simulation obtains a vertical eye opening of

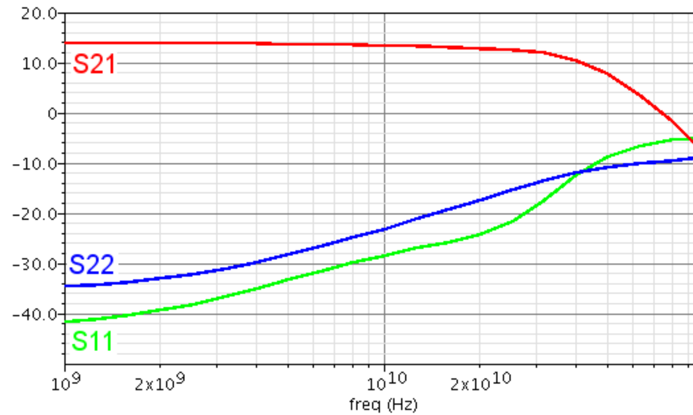


Figure 2.5: Small signal simulation results of the input buffer including parasitics and EM-models, simulated with an input and output impedance of $100\,\Omega$ *differential.*

100 mV, a horizontal eye opening of 11 ps and less than 1 ps of jitter. The PAM4 eye shows a vertical opening of 70 mV and a horizontal opening of 16 ps.

2.2.6 Comparison

The litrature on input buffers for serial electrical interconnects is limited. For this reason the input buffer is compared to transimpedance amplifiers for 40 Gb/s communication. For this comparison the TIA was simulated with a photodiode input, which results in higher gain and bandwidth compared to a $100\,\Omega$ differential input. An overview can be found in table 2.1. From the table it is clear that the proposed transimpedance amplifier achieves the highest Figure of Merit (FOM), which includes the BW, input matching and power consumption. Since this amplifier is used as an input buffer for electrical communication it will not be linked to a photodiode and silicon area is the main cost contributor. The design presented here is also the smallest (including pads) of the compared transimpedance amplifiers, which is beneficial from the point of view of cost reduction of an electrical communication system.

[2] simulated result

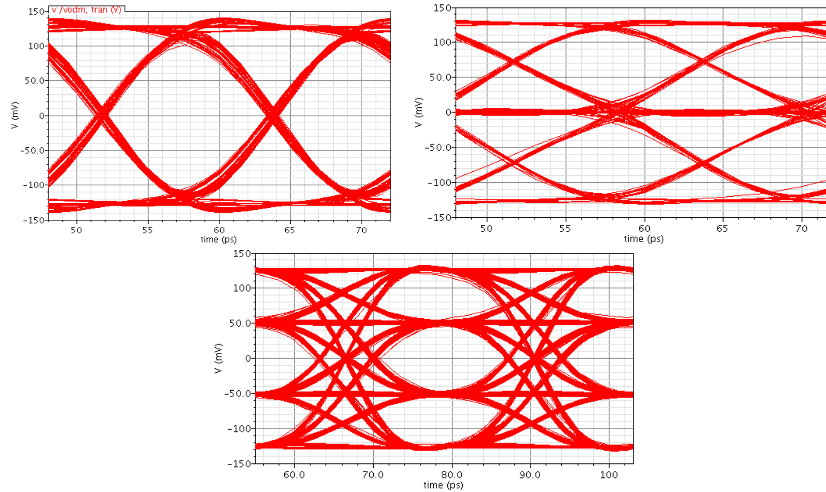


Figure 2.6: Transient simulation results of an 80 Gb/s bitstream. NRZ (top left), duobinary modulation (top right) and PAM4 (bottom) modulation are shown. The input amplitude is 60 mVpp from a $100\,\Omega$ *differential source for all signals, rise and fall times are half the bit period and the waveform is piecewise linear, chip temperature is set to 75 °C.*

2.2.7 Conclusion

In this paper a broadband input buffer is presented capable of receiving 80 Gb/s data streams in a multitude of modulation formats. The receiver has 14 dB of gain and a maximum noise figure of 8 dB in the passband. It operates from a 2.5 V power supply, consumes 80 mW and has an area of only 0.18 mm^2 including pads in a 130 nm SiGe BiCMOS technology. The circuit layout is shown in Figure 2.7.

Ref.	[6]	[7]	[8]	This work[2]
Process	SiGe f_T=200 GHz	CMOS 130 nm	SiGe f_T=220 GHz	SiGe f_T=220 GHz
Bit Rate (Gb/s)	40	40	40	80
Z_T (dBΩ)	43	50	54	54
BW of Z_T (GHz)	50	29	28	41
GD of Z_T(ps)	-	$\pm$ 8	-	$\pm$ 3
Supply (V)	5.2	1.5	3	2.5
Power (mW)	182	45.7	110	80
Area (mm^2)	0.92	0.4	0.55	0.18
FOM $(BW\Omega/P_{DC})$ (Ω/pJ)	77.4	200.7	127	248

Table 2.1: Comparison to reported 40 Gb/s transimpedance amplifiers.

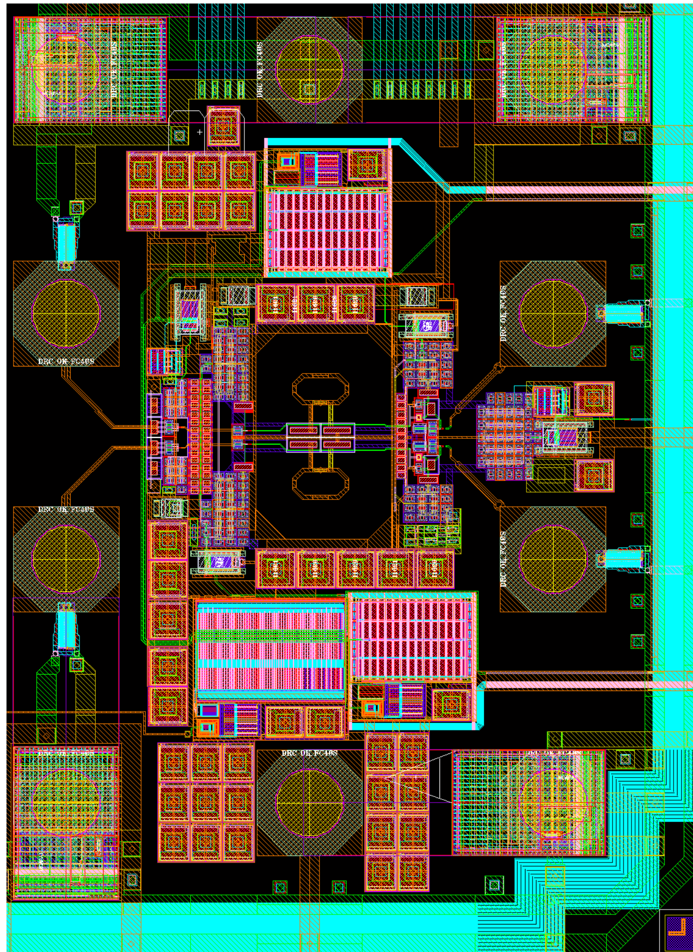


Figure 2.7: Layout input buffer with pads.

2.3 A Digitally Controlled Threshold Adjustment Circuit in a 0.13 µm SiGe BiCMOS Technology for Receiving Multilevel Signals up to 80 Gb/s

2.3.1 Introduction

Advances in the data rate of modern communication systems has lead to the need for an increase in speed of inter-chip communication. Typically, a non-return-to-zero (NRZ) signal is transmitted over a PCB transmission line (microstrip, grounded coplanar waveguide), but, with rising speed, the limited bandwidth of the channel (the transmission line) and the maximum bandwidth that can be achieved by the chip technology (indicated by the maximum transition frequency f_T) impose a limitation on the maximum data rate. To counter this limitation, multi-level signaling such as PAM4 or duobinary signaling have been used in recent papers [9–11] to demonstrate data rates up to 25 Gb/s across electrical backplanes.
Although more advanced modulation schemes require less bandwidth, the processing needed on both the transmitting and the receiving chip significantly increases. Among others, a threshold adjustment circuit (TAC) operating at high-speed is needed at the receiving end to seperate different symbol levels. This paper presents a 80 GHz TAC designed in a 0.13 µm SiGe BiCMOS technology using a 7 bit current DAC for threshold adjustment. The large bandwidth and the fine threshold control of the presented circuit allows duobinary transmission up to 80 Gb/s when signalling over a 40 GHz channel.
The TAC is presented using the following structure: in Section 2.3.2 the topology of a standard non-clocked duobinary receiver is discussed, Section 2.3.3 presents the circuit that is used to achieve a high-speed threshold adjustment, the DACs used in the control of the threshold level are shown in Section 2.3.4, Section 2.3.5 discusses the fabrication of the chip and Section 2.3.6 summarizes the results and concludes this paper.

2.3.2 Non-clocked duobinary receiver

One of the more promising modulation schemes to reduce the required channel bandwidth without adding too much chip complexity is the duobinary modulation scheme discussed in [1]. A receiver structure as well as the typical 3-level eye diagram of this modulation scheme is shown in Figure 2.8. Due to the high speeds, a fully differential channel is needed. For simplicity, only the single-ended signaling is drawn. The wideband input amplifier needs a bandwidth comparable to the channel bandwidth, for a

80 Gb/s duobinary transmission this corresponds to approximately 40 GHz. An input buffer achieving this performance is shown in Section 2.2 or [12]. After the input buffer, the signal is split and compared with both a lower and an upper threshold voltage. To do this comparison, the differential signal is first shifted using a TAC. This will position the zero level of the signal in the middle of the lower and upper eye respectively. The resulting signal is then applied to a limiting amplifier stage, which regenerates a signal with digital levels, i.e. logic high and logic low above and below the threshold respectively, necessary for the XOR gate, which will demodulate the duobinary signal.

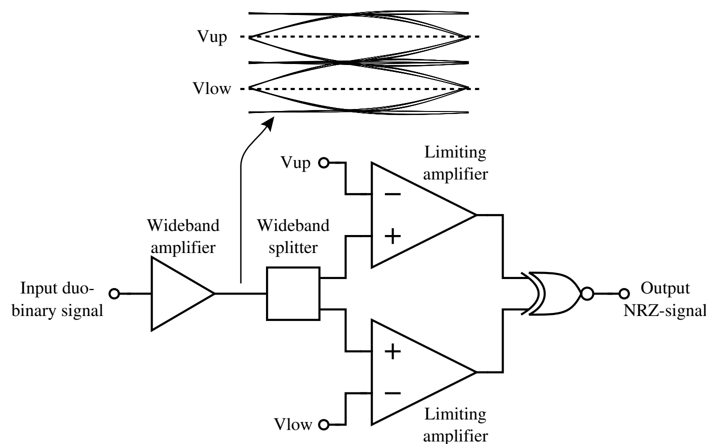


Figure 2.8: Standard non-clocked duobinary receiver as described in [1].

2.3.3 Emitter follower threshold adjustment circuit

The TAC, as shown in Figure 2.9, consists of two distinct parts: an emitter follower with threshold adjustment, and a cascoded output driver. Both circuits will be discussed in greater detail in the following sections.

2.3.3.1 Emitter follower threshold adjustment

The input of the TAC uses emitter followers, shown as transistors Q_0 and Q_1 in Figure 2.9. The transistors are biased with a constant current of 3.2 mA, which is optimized for maximal bandwidth (max f_T biasing). The use of emitter followers has two main benefits. Firstly, they provide a low output impedance, and as a result allow for a higher bandwidth when driving the capacitive input of the cascoded output stage. Secondly, the voltage

relationship between base (fixed at 2.4 V DC by the wideband input amplifier) and emitter (given the constant emitter current) is fixed. This results in an equal DC voltage at the emitters of transistors Q_0 and Q_1. To introduce a shift in DC voltage, and hence in threshold, a series resistor (R_0 and R_1) is added between the output of the emitter followers and the input of the cascoded output stage. The biasing current of the emitter follower is split into two parts, one directly connected to the emitter, one connected through this series resistor. By changing the ratio of these two current sources (varying α between 0 mA - 1.6 mA), the amount of current flowing through the resistor and hence the DC level at the input of the next stage can be controlled. By varying the DC voltage of the positive and negative input in the opposite direction, the threshold can be adjusted.

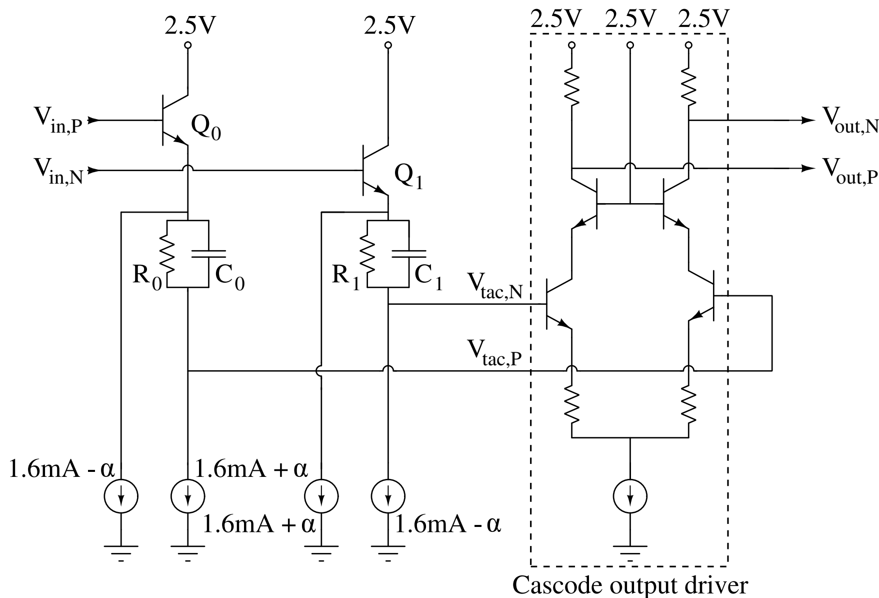


Figure 2.9: Threshold adjustment circuit consisting of emitter followers biased with a constant current and the cascade output driver. The threshold level is adjusted by means of α ranging from 0 mA - 1.3 mA.

Furthermore, capacitors C_0 and C_1 are added in parallel with resistors R_0 and R_1 respectively in Figure 2.9. This provides a low impedance path at higher frequencies between the output of the emitter followers and the input of the cascode and hence, increases the bandwidth.
The capacitors used in this circuit are metal-insulator-metal (MIM) capacitors with their bottom plate connected to the emitter of the emitter followers and their top plate connected to the input of the cascode output driver. This

reduces the parasitic capacitance at the input of the cascoded amplifier and increases the bandwidth. The use of this topology allows to shift signals with a bandwidth of more than 80 GHz as shown in the post layout simulation results of Figure 2.10. The variation on the bandwidth is less than 5% across the whole shifting range.
The shifting resistors R_0 and R_1 have a value of $80\,\Omega$ in this design. The following trade-off exists: larger resistor values give a larger maximum threshold adjustment but a smaller resolution and require a higher bypass capacitance while smaller resistor values mainly deteriorate the maximum threshold variation. Parallel with these $80\,\Omega$ resistors a 1 pF MIM capacitor was implemented.

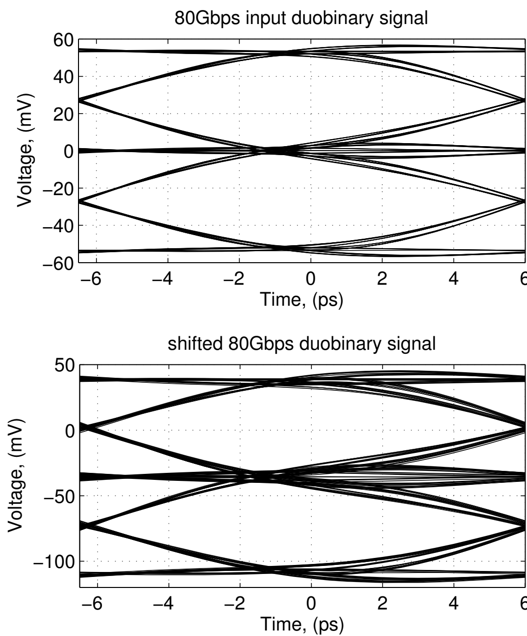


Figure 2.10: Input eye diagram of a 80 Gb/s duobinary input signal together with the eye diagram of the shifted version at the output of the threshold adjustment circuit.

The current mirrors providing the current to the emitter followers are cascoded current mirrors. This reduces the effect of the varying collector voltages of the mirror transistor. This allows us to have a current resolution of less than 5 µA resulting in a voltage shifting resolution of less than 1 mV.

2.3.3.2 Cascoded output driver

The switching output driver of which the threshold is adjusted, consists of a degenerated differential cascode as shown in Figure 2.9. By using a cascode, the miller capacitance at the input can be diminished [13] and the maximally allowed voltage swing at the output is higher since the peak voltage at the collector of a common base stage can go up to BV_{CBO} [14, 15]. Feedback by means of resistive degeneration is applied to increase the bandwidth of the circuit by trading in gain. Furthermore, it makes the cascode less sensitive to differences in DC-voltage at the input which limits the accuracy that is needed for the TAC. The biasing of the cascode is such that the transistors achieve their maximum f_T. The load resistor is a $80\,\Omega$ poly resistor chosen to allow enough bandwidth when loaded with the input capacitance of the next stage. The tail current is 5 mA.
In Figure 2.10, the output eye diagram shows a zero-volt level that corresponds to the eye crossing of the upper duobinary eye. By adding a limiting amplifier to this output, the wanted NRZ-signal representing the upper half of the duobinary signal is obtained.

2.3.3.3 Bandwidth and linearity

In Figure 2.11, the gain and the input referred 1-dB compression point are simulated as a function of frequency. The 3-dB bandwidth of the circuit is above 80 GHz, to ensure that the level shifting stage will not limit the bandwidth in a 80 Gb/s data system. The simulated input referred 1-dB compression point is around -4 dBm (400 mVpp).

The 1-dB compression point was simulated by driving power from a $50\,\Omega$ input buffer connected in front of the TAC and measuring the input power and the output power of the TAC. Since the linearity of the input buffer is greater than the linearity of the TAC, this is a valid measure for the linearity. At higher frequencies it is no longer possible to include a sufficient number of harmonics in the simulation, which leads to a lower simulated 1-dB compression point, as can be noticed from the two rightmost 1-dB compression points in Figure 2.11. Verification of the linearity at high frequencies was subsequently done by checking the in- and output eye diagram at different input power levels.

2.3.4 Digitally controlled current DACs

The circuit used to generate the currents for the emitter followers is shown in Figure 2.12, both currents are derived from one variable current source

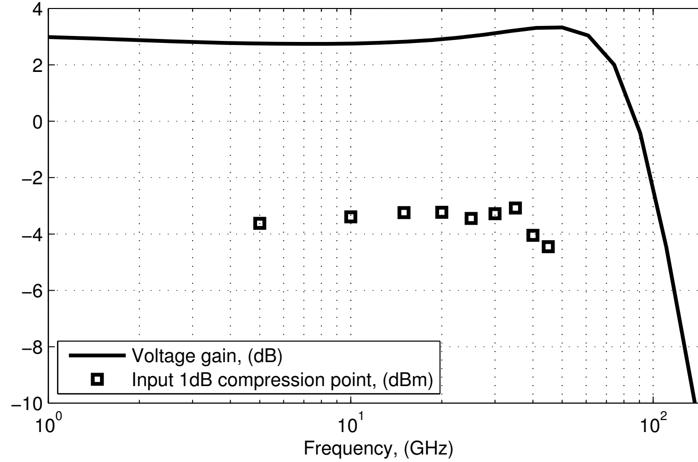


Figure 2.11: Post layout simulation results for the voltage gain and input referred 1-dB compression point of the TAC as function of frequency.

I_{var}. This source is implemented as a traditional binary weighted current mirror with switchable gates to control the total current. The variable current is mirrored and amplified five times. This amplification allows to reduce the current in the DAC as well as the area consumed by the DAC. However, care has to be taken to meet the matching requirements for the 7 bit DAC resolution used in this design.
In a second stage, the amplified variable current is added and subtracted from a constant current source I_0. Using this technique the current through the emitter followers will always be twice I_0. The last stage of the current generator implements a switch which allows digital control of the shifting direction.

Measurements on the fabricated chip showed that the current switching circuit has a minimum of 263 µA and a maximum of 1.26 mA, adjustable in steps of 4 µA. Taking into account the $80\,\Omega$ resistor that converts this current into the level shifting voltage and the times two multiplication in the emitter follower current mirrors, we can shift the level 160 mV up or down in steps of 0.64 mV. The current running through the emitter followers is constant around 3 mA.

2.3.5 Fabrication and further Work

A chip implementing the discussed threshold adjustment circuit has been fabricated in a 0.13 µm SiGe BiCMOS technology with an f_T beyond

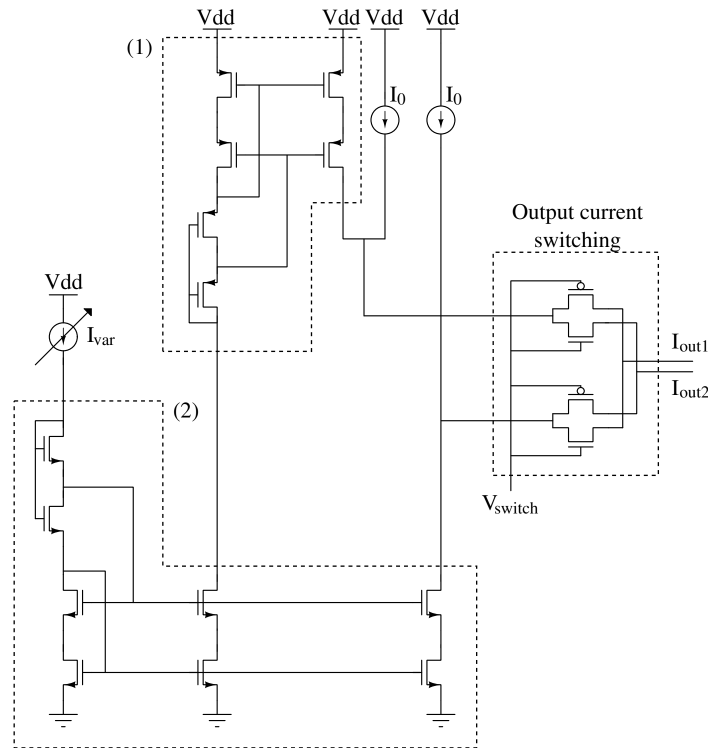


Figure 2.12: Circuit used to generate complementary output currents I_{out1} and I_{out2} from the DAC current I_{var}. Sooch cascode current mirrors, subcircuits (1) and (2), are used to get accurate copies of the DAC current. The digital signal V_{switch} interchanges the output currents.

200 GHz. The part of the chip with the discussed circuit is shown in the micrograph of Figure 2.13. Although the proposed circuit can not be tested on its own, measurements show that an eye at the input can be shifted over the complete input range.

Further testing of the chip implementing the discussed circuit, includes the reception and demodulation of a 80+ Gb/s duobinary signal over a 40+ GHz channel. With improvement of the channel bandwidth even higher data rates should be attainable since the bandwidth of the TAC is as high as 80 GHz.

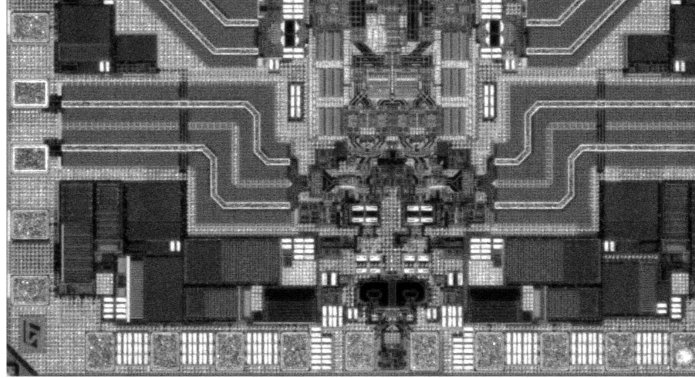


Figure 2.13: Die micrograph of the TAC circuit.

2.3.6 Conclusion and results

In this section a broadband threshold adjustment circuit capable of level shifting a duobinary data stream with a bandwidth up to 80 GHz is presented. The threshold adjustment circuit can shift a differential input signal up or down by 160 mV with a resolution of 0.64 mV. It operates from a single 2.5 V power supply, consumes only 35 mW in a 130 µm SiGe BiCMOS technology and has a relatively high input referred 1-dB compression point of approximately -4 dBm. Using this circuit it is possible to demodulate a multitude of modulation schemes (e.g. duobinary, PAM4) at high bandwidths allowing to achieve higher data rates over channels with limited bandwidth while keeping the added circuit complexity low.

2.4 Additional high-speed subblocks

2.4.1 Combining the level shifters into the LSLA

The duobinary front end, as explained in Section 2.1, consists of a TIA input buffer and a level shifting limiting amplifier (LSLA). The TIA has sufficient bandwidth to receive a duobinary signal up to roughly 100 Gb/s on chip and distribute it to the LSLAs, which transfer the signal to the sampling stage of the deserializer-XOR structure.
The LSLA shown in Figure 2.14 consists of two traditional CML gain stages, two level shifting stages, as explained in Section 2.3, a buffer connected to the sampling stage and a buffer connected to an output driver to be able to display the upper or lower eye off-chip.

The first stage of the LSLA is used to shift the data up (or down) with a 8 bit digital control word (7 bit level control and 1 sign bit) to retrieve the upper (or lower) eye from the duobinary stream. Due to the unbalance in the recovered eye (ideally 25% top and 75% bottom mark densities) the crossing will not be in the middle of the eye anymore after amplification. To overcome this, a second level shifting stage is added which provides coarse control of the eye crossing. This technique is the subject of a filed patent application (EP14161772.0).

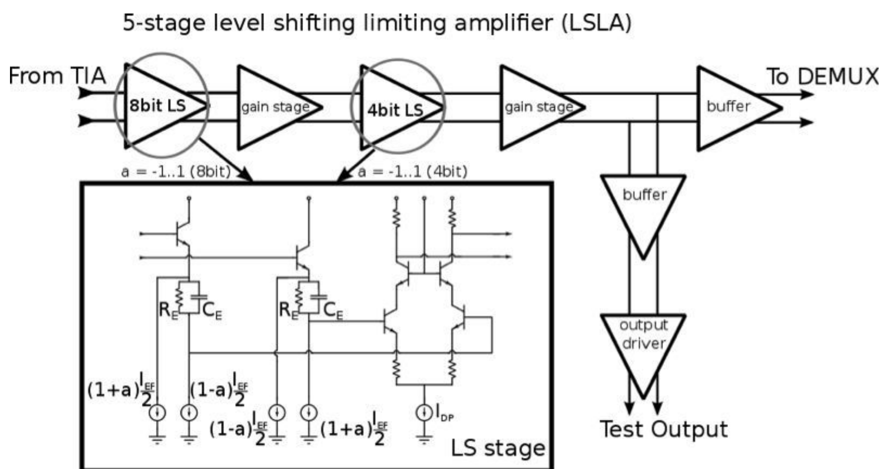


Figure 2.14: Block diagram of the LSLA, including the level shifting stages.

2.4.2 Sampling stages

The high-speed sampling stage consists of two CML latch stages with anti-phase differential half rate clock inputs. When the clock is high, the sampling stage acquires the data and makes the decision on whether the input is logic high or logic low. When the clock is low, the data is regenerated to digital CML levels. This stage is succeeded by another latch stage which samples the regenerated data to make sure the output of the sampling stage is a digital half rate CML signal.
This is illustrated in the block diagram of Figure 2.15, underneath the block diagram the schematic of the latch is shown. The load resistors in each latch stage are 55 Ω, connected to the 2.5 V supply. This realizes about 400 mV differential swing for the XOR gate with a bias current around 4 mA.

In total, there are four full speed sampling stages and four half rate sampling stages on the receiver die, as shown in Figure 2.1. The clock is distributed in

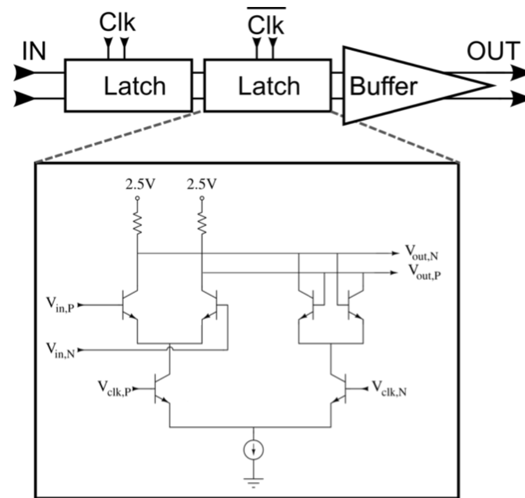


Figure 2.15: Block diagram of the high-speed sampling stage, including latch schematic.

such a way that two full speed sampling stages sample at the rising edge and the other two stages sample at the falling edge, this effectively divides the bit rate by two. Each half bit rate signal is then processed by the XOR gates to decode the duobinary data. After decoding, the half rate sampling stages deserialize the stream in four quarter rate re-timed differential outputs.

2.4.3 Half rate XOR gate

In this section the design of a high-speed XOR gate, demonstrated to work up to 42 Gb/s is shown. The design is based on a low voltage topology using switched emitter followers, it was designed in close collaboration with my colleague Zhisheng Li.

After de-multiplexing the two full rate input data streams into four half rate data streams, the XOR function is applied. The schematic of the XOR gate is depicted in Figure 2.16. Since the output of the XOR gate will be further demultiplexed to four quarter rate output data signals the delay of the XOR gate including data buffers between the XOR gate and the quarter rate sampling stages is important, to ensure the correct timing for the quarter rate sampling stages. This delay is simulated to be 20 ps.

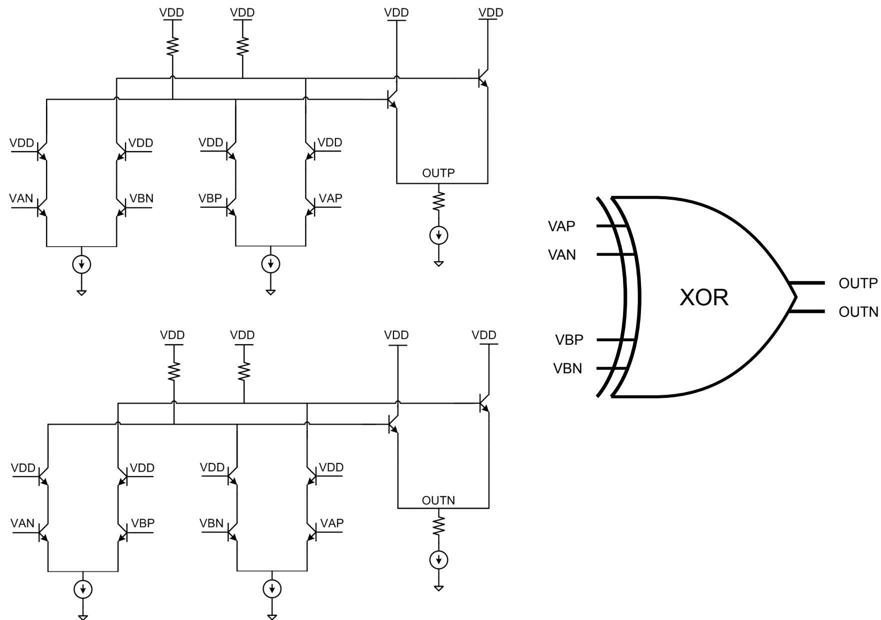


Figure 2.16: Schematic of the XOR gate.

2.4.4 Quarter rate output buffers

The four NRZ data outputs of the DEMUX are buffered by the output stage shown in Figure 2.17, using a $50\,\Omega$ CML output stage. An emitter follower with a DC-block is placed in front of the amplifier to eliminate any DC offset. The internal buffers of the two DEMUX input signals coming from the receiver front end are not shown here.

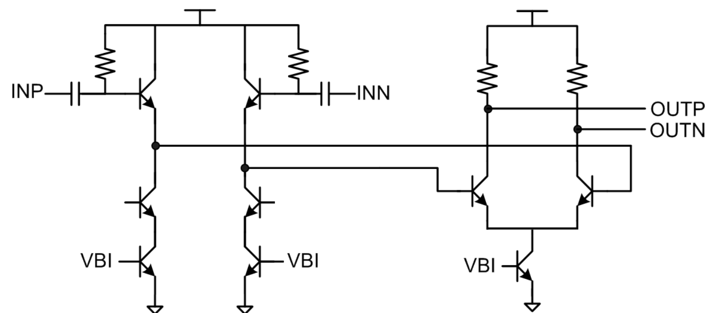


Figure 2.17: Schematic of the data output buffer.

2.4.5 Clock chain

The clock chain was designed by my colleague Zhisheng Li. To cover the whole data range (10 to 100 Gb/s), a $10\,\mathrm{ps}$ programmable delay range is required. The schematic of a programmable unit-delay buffer is shown in Figure 2.18. When only $EN0$ is enabled, it has $0\,\mathrm{ps}$ relative delay. When only $EN1$ is enabled, the delay is around $1.9\,\mathrm{ps}$. While enabling both $EN0$ and $EN1$ results in a delay of around $1.2\,\mathrm{ps}$. By cascading six of these unit-delay buffers, the maximum delay is $11.4\,\mathrm{ps}$.

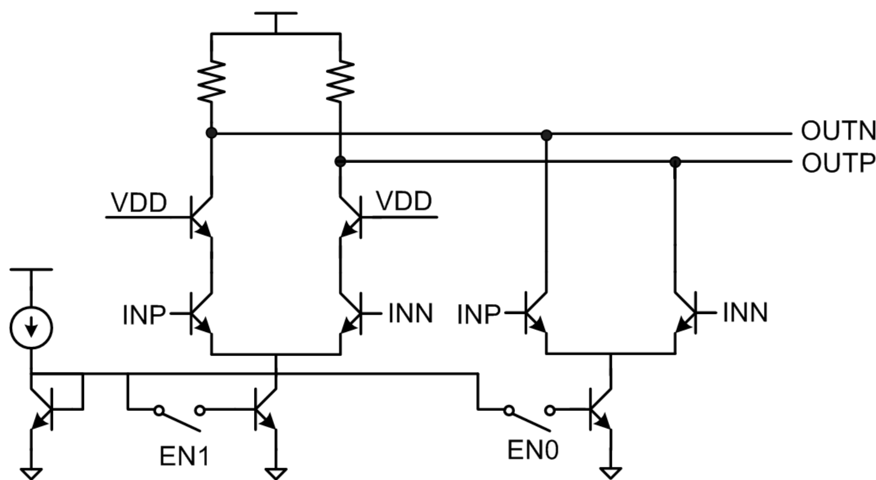


Figure 2.18: Schematic of the programmable delay cell.

The final quarter rate outputs are implemented using four sampling stages working at a half rate clock frequency. The half rate clock signals should have quadrature phases, which are generated by the divider-by-2 circuit (DIV2). The core circuit of the DIV2 block is the standard Master-Slave topology, as shown in Figure 2.19, which is the same as the DIV2 block used in the MUX at the transmitter. Both the input and output signals are buffered. The D-latch of the DIV2 circuit is based on a standard Gilbert cell.

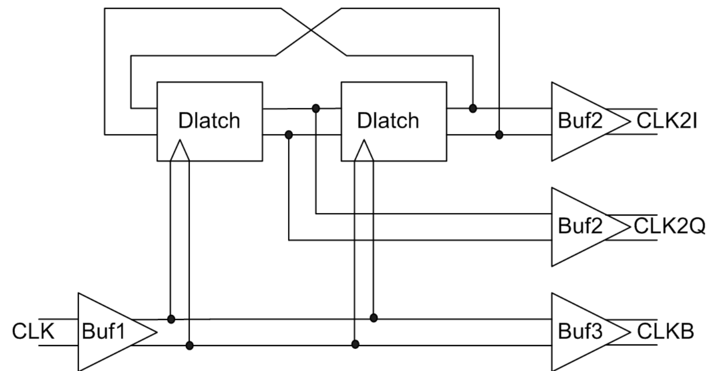


Figure 2.19: The topology of the divide-by-2 circuit

2.5 Power consumption overview

The most important building blocks were discussed in conference publications and incorporated in this chapter. Total power consumption at the receiver side adds up to 1.4 W for a 100 Gb/s data transmission resulting in 14 pJ/bit including SERDES. In Figure 2.20 it is shown that 65% of the power is consumed in the XOR/DEMUX circuit and 35% in the receiver front end . The power consumed in the front end is about 500mW and is

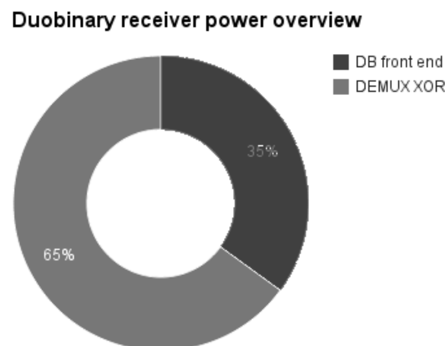


Figure 2.20: Overview of power division between duobinary front end and DEMUX/XOR.

mainly consumed by the level shifting limiting amplifier (LSLA), a more in detail overview of the power consumption in the duobinary front end is given in Figure 2.21a. The largest amount of power dissipation occurs at the DEMUX/XOR of which a detailed overview is depicted in Figure 2.21b

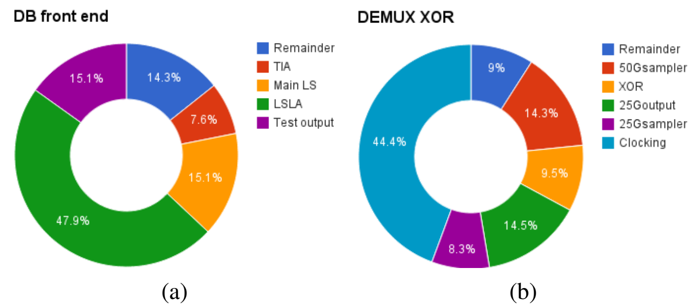


Figure 2.21: Overview of power consumption division between duobinary front end (a) and DEMUX/XOR (b).

2.6 Conclusions

In this chapter an overview was given of the design of the duobinary receiver chip. A die photograph of the designed duobinary receiver is shown in Figure 2.22. The chip measures 1.9 mm by 2.6 mm and was flip-chipped directly on the test boards. This receiver chip allows to demodulate duobinary transmissions up to 100 Gb/s with 14 pJ/bit. On-chip demulitiplexing is integrated which delivers four quarter-rate NRZ streams corresponding to the received duobinary signal.

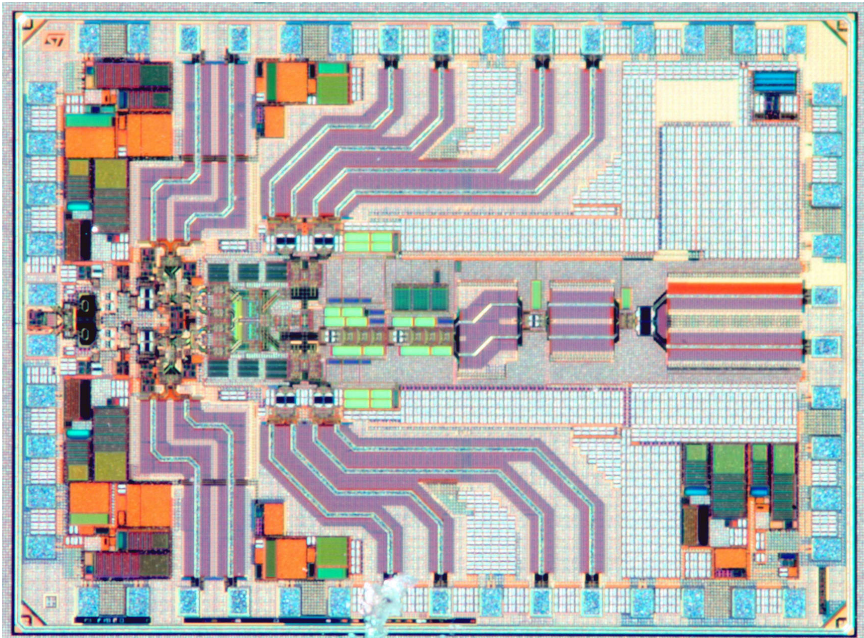


Figure 2.22: Die photograph of the designed duobinary receiver.

References

[1] J.H. Sinsky, M. Duelk, and A. Adamiecki. High-speed electrical backplane transmission using duobinary signaling. *IEEE Transactions on Microwave Theory and Techniques*, 53(1):152–160, 2005.

[2] T. Moselhy. *Field Solver Technologies for Variation-Aware Interconnect Parasitic Extraction*. PhD Dissertation, Massachusetts Institute of Technology, 2010.

[3] D. Cascella, G. Avitabile, F. Cannone, and G. Coviello. A 2-gs/s 0.35μm SiGe track-and-hold amplifier with 7-ghz analog bandwidth using a novel input buffer. In *2011 18th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 113–116. IEEE, 2011.

[4] T. Dickson, R. Beerkens, and SP. Voinigescu. A 2.5-v 45 Gb/s decision circuit using SiGe bicmos logic. *IEEE Journal of Solid-State Circuits*, 40(4):994–1003, 2005.

[5] A. Jentzsch. *Elektromagnetische Eigenschaften von Flip-Chip-Übergängen im Millimeterwellenbereich*. PhD Dissertation, TU Berlin, 2002.

[6] J. Weiner, A. Leven, V. Houtsma, Y. Baeyens, YK. Chen, P. Paschke, Y. Yang, J. Frackoviak, WJ. Sung, A. Tate, et al. Sige differential transimpedance amplifier with 50-ghz bandwidth. *IEEE Journal of Solid-State Circuits*, 38(9):1512–1517, 2003.

[7] J. Kim and JF. Buckwalter. Bandwidth enhancement with low group-delay variation for a 40 Gb/s transimpedance amplifier. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(8):1964–1972, 2010.

[8] SB. Amid, C. Plett, and P. Schvan. Fully differential, 40 Gb/s regulated cascode transimpedance amplifier in 0.13 μm SiGe bicmos technology. In *2010 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pages 33–36, 2010.

[9] D. Correia, V. Shah, CP. Chua, and P. Amleshi. Performance comparison of different encoding schemes in backplane channel at 25+ Gb/s. In *2013 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pages 306–311, 2013.

[10] J. Lee, M.S. Chen, and H.D. Wang. Design and comparison of three 20 Gb/s backplane transceivers for duobinary, pam4, and nrz data. *IEEE Journal of Solid-State Circuits*, 43(9):2120–2133, 2008.

[11] A. Adamiecki, M. Duelk, and J.H. Sinsky. 25 Gb/s electrical duobinary transmission over fr-4 backplanes. *Electronics Letters*, 41(14):826–827, 2005.

[12] T. De Keulenaer, Y. Ban, Z. Li, and J. Bauwelinck. Design of a 80 Gb/s SiGe bicmos fully differential input buffer for serial electrical communication. In *2012 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 237–239. IEEE, 2012.

[13] E. Säckinger. *Broadband circuits for optical fiber communication*. John Wiley & Sons, 2005.

[14] H. Li, H.M. Rein, T. Suttorp, and J. Bock. Fully integrated SiGe vcos with powerful output buffer for 77-ghz automotive radar systems and applications around 100 ghz. *IEEE Journal of Solid-State Circuits*, 39(10):1650–1658, 2004.

[15] M. Rickelt, H.M. Rein, and E. Rose. Influence of impact-ionization-induced instabilities on the maximum usable output voltage of si-bipolar transistors. *IEEE Transactions on Electron Devices*, 48(4):774–783, 2001.

3

Measurements of millimeter wave test structures for high-speed chip testing

Timothy De Keulenaer, Yu Ban, Guy Torfs,
Stefaan Sercu, Jan De Geest and Johan Bauwelinck

Based on the article presented at the IEEE Workshop on Signal and Power Integrity (SPI 2014) extended with comments on flip chip mounting, application specific board design and footprint design.

The authors would like to thank the Agency for Innovation by Science and Technology in Flanders (IWT) and the INTEC EM group for their 2D impedance calculator.

⋆ ⋆ ⋆

THIS chapter presents the frequency domain characterization of very high-bandwidth connectorized traces and a millimeter wave rat-race coupler. The connectorized differential grounded coplanar waveguide traces, essential for the testability of high-speed integrated circuits, have a mea-

sured smooth frequency response up to 67 GHz, which indicates correct connector footprint and transmission line design. The differential traces narrow down to the bondpad pitch of 150 μm allowing direct flip chip connections. In this way enabling the testing of millimeter wave integrated circuits without the need for probing. Furthermore, a 50 GHz rat-race coupler was fabricated to generate a differential clock from a single-ended clock source.

3.1 Introduction

At high data rates, inter-chip electrical communication over standard traces becomes challenging due to excessive frequency dependent channel attenuation causing large amounts of inter-symbol interference (ISI). Combined with impedance mismatch, a very challenging environment for high-speed and high-bandwidth communication is created. When testing high-speed communication chips, care must be taken in the design of the test board to ensure measurements reflect the chip performance, and not the connectorized test board.

High-bandwidth integrated circuit designs face a variety of technical difficulties. First of all, there are the functional and matching requirements, which in the presence of layout parasitics, ESD protection mechanisms and process limitations, can be hard to reach. However, after fabrication, a second problem arises: testing a chip in real life circumstances requires that it is mounted on a board, using either bondwires, a flip chip process or some other kind of packaging.

Typically, direct chip-on-board flip chip assembly (without interposer) adds the least amount of parasitic inductance and capacitance, making it the preferred method for high-speed/high-bandwidth chip-to-board interconnects with bandwidths above 50 GHz.

However, to be able to flip a chip directly on a board, the pitch of the traces on the board and the bondpads on the chip need to align. This clarifies the need of a fine pitched, differential grounded coplanar waveguide (GCPW) structure such as shown in Figure 3.1. On the other side of this transmission line structure, single-ended connectors are used to connect the test board to high-speed test and measurement equipment like a sampling scope, vector network analyzer or BER tester.

Furthermore, high-speed chips typically need a differential clock to synchronize data to the test and measurement equipment, unless an on-chip CDR or balancing circuit is available, however this adds complexity and consumes scarce chip real estate. Providing a differential clock at high frequencies requires a single-ended to differential converter as most millimeter

wave frequency generators only provide a single-ended output. To convert this single-ended output to a differential clock a modified 50 GHz rat-race coupler was designed and measured. Section 3.2 describes the design as well as the sizing of a GCPW transmission line and a rat-race coupler. In Section 3.3, different test structures are defined and the measurements of these test structures and the rat-race coupler are discussed. The chapter will end with a conclusion on how to select the right connectors and how to design transmission lines for testing mounted high-speed chips.

3.2 Design of millimeter wave test structures

A high-speed test board typically consists of coupled transmission lines starting at the chip, routed to connectors, together with some lower speed control signals and power and ground connections. Most challenges are found in the coupled transmission lines starting from the IC bondpad pitch and tapering out to a reasonable dimension to reduce loss and the effect of manufacturing tolerances. Further on, the coupled traces are split into 2 single transmission lines and connected out using a 50 Ω connector. In Figure 3.1 a close up is shown of a typical high-speed test interconnection.

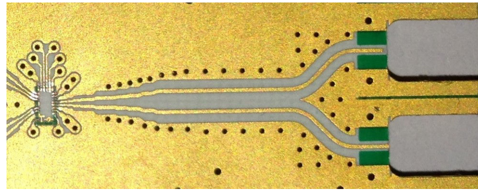


Figure 3.1: Typical high-speed test interconnection, traces taper out from the bondpad pitch to the connector footprint. In this case a miniSMP footprint is shown.

3.2.1 Coupled GCPW

To calculate the impedance of a coupled grounded coplanar waveguide structure, a 2D impedance calculator developed by the Ghent University INTEC EM group was used [1]. This tool allows one to draw any number of dielectrics and conductors and calculates the differential and common mode impedance of 2 single conductors given the dielectric constants, metal thickness and dielectric thickness, which can be found in the material datasheet or in the PCB manufacturer documentation. A structure as shown

at the top of Figure 3.2 is drawn and calculated.
The differential impedance is designed to be $100\,\Omega$. However, to connect the chip directly on the traces, a pitch of $150\,\mu\text{m}$ is required. To manage this within manufacturing possibilities the differential impedance will deviate from $100\,\Omega$ as the gap (G) and the clearance (C) are equal to the minimum manufacturable clearance of $70\,\mu\text{m}$ and traces (W) are sized to be $80\,\mu\text{m}$, which for the $221\,\mu\text{m}$ RO4003C with $30\,\mu\text{m}$ of plated top metal results in a differential impedance of $120\,\Omega$. Further on, the traces are tapered out by a concatenation of tapers to realize the $100\,\Omega$ differential impedance and to reduce the influence of manufacturing tolerances. It should be noted that the common mode impedance varies across the concatenation of tapers.

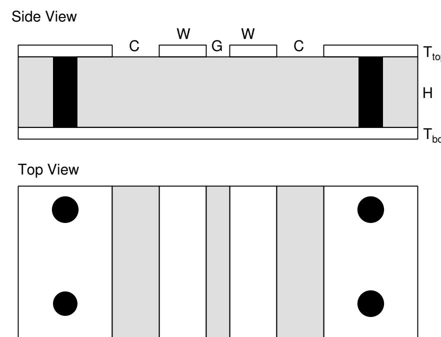


Figure 3.2: Side and top view of a coupled grounded coplanar waveguide with design parameters: W (width), G (gap), C (clearance), H (height), via size, via spacing and distance between vias and traces. T_{top} and T_{bot} are respectivly the top metal thickness including plating and the bottom metal thickness.

As explained in [2] the via spacing, via distance towards the edge of the ground plane and the via diameter have an influence on the bandwidth of the trace. Typically, reducing these distances will improve the bandwidth. Furthermore, the designed test structures use a double via row as advised in the Rosenberger $1.85\,\text{mm}$ reference footprint [3]. The length of the different tapers is always chosen equal to the distance between the vias for layout convenience.

3.2.2 Coupled to uncoupled GCPW splitter

To connect the coupled transmission lines to measurement equipment, it needs to be split into two single-ended transmission lines so that a connector can be mounted at the end of these transmission lines. The dimensions of the coupled and uncoupled coplanar waveguide are used as a starting point of the coupled to uncoupled GCPW splitter. These are connected together

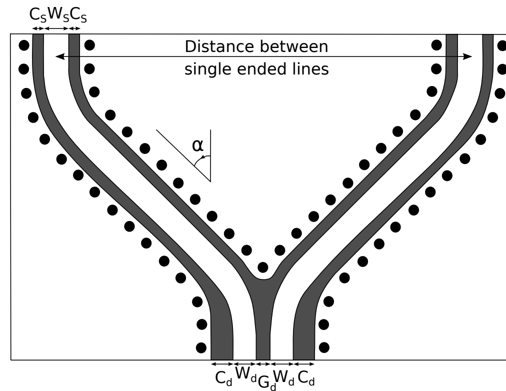


Figure 3.3: Coupled to uncoupled GCPW splitter on pcb board. C_d, W_d and G_d are the parameters defining the differential impedance and C_s and W_s define the impedance of the uncoupled lines.

in a smooth way to minimize the deviation from a $100\,\Omega$ differential line and, hence, maximize the bandwidth. As parameters we took the angle (α) and the distance between the single-ended lines as shown in figure 3.3. Optimizing α with a fixed distance led to a maximum bandwidth at an angle of around 60°, which resulted in a single-ended to differential conversion with a simulated bandwidth more than $100\,\mathrm{GHz}$.

3.2.3 Connector footprint

Test structures were developed to test 2 types of screw-on connectors and a few different footprints for the edge mount Rosenberger mini-SMP connector (18S203-40M L5). The Southwest Microwave $2.4\,\mathrm{mm}$ end launch connector (1492-03A-5) and the Rosenberger $1.85\,\mathrm{mm}$ angle launch connector (08K80A-40M) are both reusable but expensive, at about 75 euro for the Southwest and 280 euro for the Rosenberger connector. Compared to only 5 euro for the Rosenberger mini-SMP connector, which is not reusable since it is soldered to the board.
The footprint design of a connector has a significant influence on the bandwidth of the system. The footprints of all connectors were optimized in CST Microwave Studio (CST MWS) starting from the footprint advised by the manufacturer [3, 4] or, if the simulation model was not available, a tailored footprint was requested from the manufacturer which was the case for the mini-SMP connector. A lab build connector model for the $2.4\,\mathrm{mm}$ end launch connector is shown in Figure 3.4. This model was used to optimize the footprint based on Time-Domain Reflectometry (TDR) analysis

as shown in Figure 3.4d. In a TDR graph the impedance at a certain point in time reflects the impedance at a certain distance from the launch. If the impedance deviates from $50\,\Omega$, it can be increased by making the section more narrow or by increasing the gap to ground or the other way around, if the impedance needs to be decreased (assuming the substrate is fixed). The corresponding return loss is shown in Figure 3.4c. Since it is below -10 dB up to 100 GHz this footprint was selected on the test board. For the angled Rosenberger connector, CST MWS simulations were carried out by Jan De Geest from FCI because the model could not be shared due to an NDA between FCI and Rosenberger.

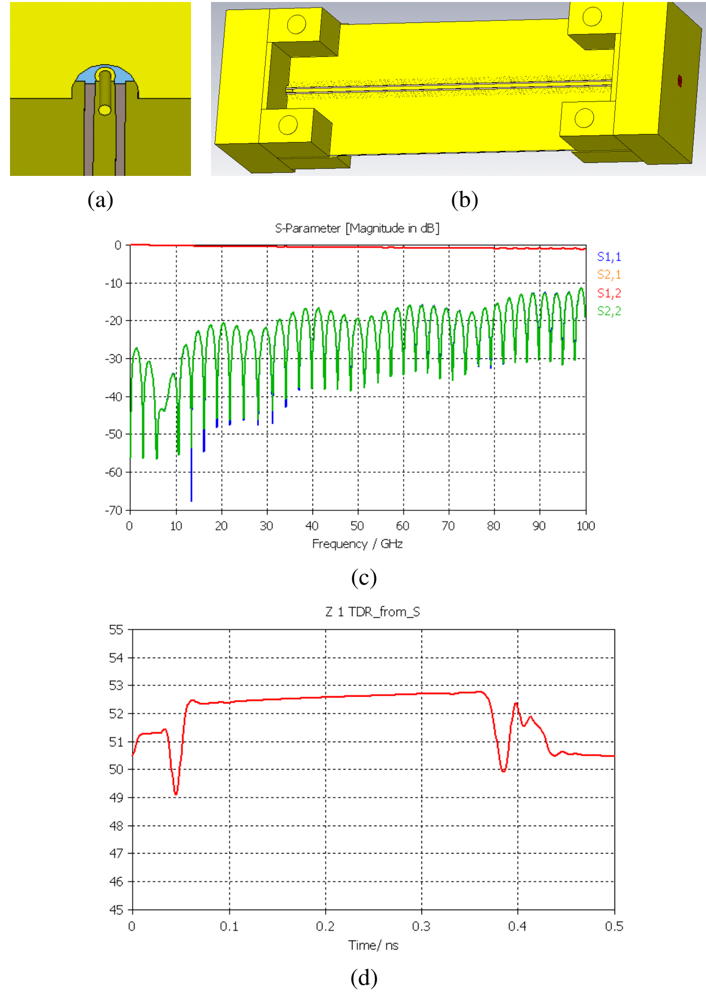


Figure 3.4: Footprint design for the 2.4 mm end launch SouthWest connector (a), simulation in CST MWS using transmission line approach (b). The resulting S-parameters (c) and TDR (d) showing a good impedance matching.

3.2.4 Rat-race coupler

Fabricating traditional rat-race couplers for high frequencies becomes a great challenge as a $70.7\,\Omega$ circle with a radius of $R = \frac{3\lambda}{4\pi}$ is needed [5]. However, as shown in Figure 3.5, it is possible to increase the radius of the circle and keep the same differential relationship at the outputs by adding $\frac{\lambda}{2}$ to the short paths and $\frac{3\lambda}{2}$ to the long path. This results in a radius of $R = \frac{9\lambda}{4\pi}$. As a consequence, this rat-race coupler also works at $\frac{9\lambda_1}{4\pi} = \frac{3\lambda_2}{4\pi}$,

one third of the design frequency.

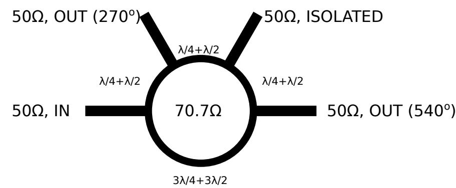


Figure 3.5: Topology of the proposed rat-race coupler, half wavelength segments are added to increase the size for improved manufacturability.

3.3 Measurements of high-speed test boards

The design methods described above were put into practice on a board design with various test structures fabricated by Wrekin Circuits[1]. On the test boards, multiple structures were placed together to test the different connectors and material losses. The boards were produced on a gold plated Rogers RO4003C substrate to resemble a test board for flip chip or wire bond assembly.
All measurements were done with a 4 port 67 GHz Agilent PNA-X.

3.3.1 Material characteristics

The different test structures were fabricated on a 221 µm Rogers RO4003C substrate cladded with 17 µm LoPro foil (LoPro) as well as on a 203 µm Rogers RO4003C substrate cladded with 17 µm electrodeposited copper foil (NoLoPro). As the interest goes to bandwidths above 50 GHz, the loss is compared at 50 GHz and shown in Figure 3.6 and Figure 3.7. It is clear that for a single-ended transmission line the loss at 50 GHz on the LoPro material is around 1.1 dB per centimeter and is above 1.4 dB per centimeter on the NoLoPro material. A summary is given in Table 3.1.
The impedance of the line is mainly determined by the dielectric constant (ϵ_R) and the effective dimensions of the trace. Figure 3.8 illustrates that the TDR impedances are about 55 Ω for the LoPro and 52 Ω for the NoLoPro material. Careful measurement of the lines show a reduction of approximately 10% in width compared to the designed value, within manufacturer tolerances but resulting in this impedance rise. Multiple samples of each

[1] Wrekin Circuits, World Class PCBs to the Electronics Industry. http://www.wrekin-circuits.co.uk

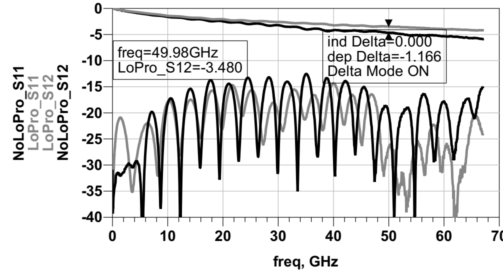


Figure 3.6: Measurements of a 26 mm *single-ended trace with* 1.85 mm *connectors.*

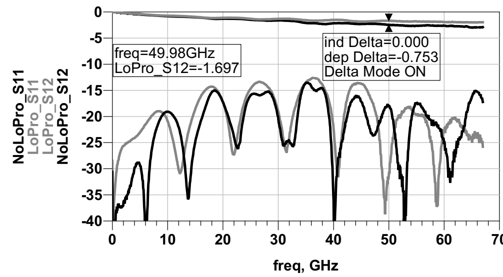


Figure 3.7: Measurements of a 10 mm *single-ended trace with* 1.85 mm *connectors.*

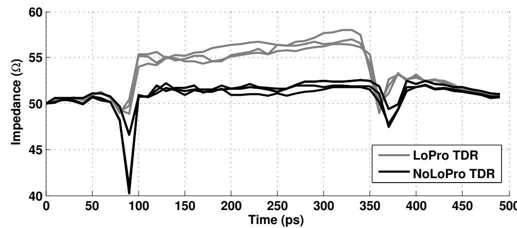


Figure 3.8: TDR measurements of 26 mm *traces, showing a capacitive drop at the connector footprint.*

trace were measured and depicted in Figure 3.8. From the figure, little variation in impedance over the different samples can be seen. The sizing of the designs fabricated on LoPro and NoLoPro material are the same, however, the NoLoPro material is a bit thinner and has a slightly higher ϵ_R which results in the lower impedances measured. To characterize the losses of the material up to 67 GHz, 1.85mm connectors were connected to 2 different line lengths, 26 mm and 10 mm. This results in the losses shown in Table 3.1.

loss at 50 GHz (dB)	10 mm	26 mm	loss per cm
LoPro	1.70	3.48	1.1
NoLoPro	2.45	4.65	1.4

loss at 30 GHz (dB)	10 mm	26 mm	loss per cm
LoPro	1.32	2.60	0.8
NoLoPro	1.90	3.22	0.83

Table 3.1: Substrate loss at 50 GHz *and at* 30 GHz *calculated from a* 26 mm *trace and a* 10 mm *trace using* 1.85 mm *connectors.*

loss at 30 GHz (dB)	30.4 mm trace
LoPro	3.5
NoLoPro	4.5

Table 3.2: Loss at 30 GHz *of a* 30.4 mm *trace on LoPro and NoLoPro substrate using* 2.4 mm *connectors.*

3.3.2 Connector characteristics

The screw-on connectors are Southwest Microwave 2.4 mm end launch connectors and Rosenberger 1.85 mm angle mount connectors. Both connectors show very clean results. The insertion loss of the 1.85 mm connector is about 0.3 dB on LoPro material and 0.5 dB on NoLoPro material at 50 GHz, as can be calculated from Table 3.1. At 30 GHz, the losses of both connectors are comparable.
The loss of the 2.4 mm connector is about 0.5 dB on LoPro material and 1 dB on NoLoPro material at 30 GHz as can be calculated from Table 3.2. The extra loss on the NoLoPro boards is likely to be caused by a larger capacitive drop at the connector footprint, as is illustrated in Figure 3.8 for a 26 mm trace.

3.3.3 Bandwidth of differential lines

The coupled tapered traces measured in this chapter consisted of a multi-stage linear taper to go from the IC bondpad pitch of 150 µm to a more comfortable pitch (to reduce the influence of manufacturing tolerances) as is shown in Figure 3.9. In simulation, an ideal taper was also analyzed by optimizing the taper shape using the tool mentioned in Section 3.2.1. The simulations did not show a difference between the multi stage linear taper and the ideal taper below 100 GHz, the maximum simulation frequency, because the length of the individual linear tapers is small compared to the

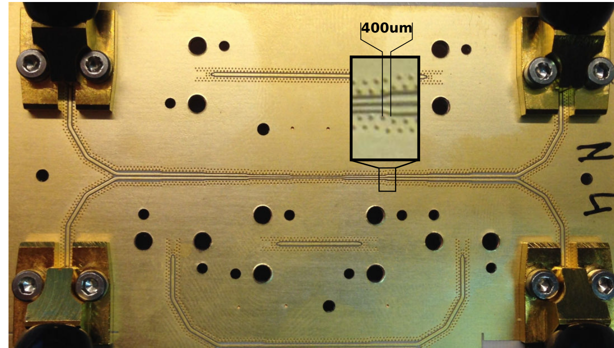


Figure 3.9: Photograph of the measured coupled trace tapering to a pitch of 150 µm. *Measurement results are shown in figure 3.10.*

frequency of operation and only shows a minor deviation from the ideal 100 Ω differential impedance.
The loss measured at 50 GHz of the 7.5 cm tapered trace is almost 11 dB, however the trace loss is smooth and thus can easily be compensated by an equalizer.

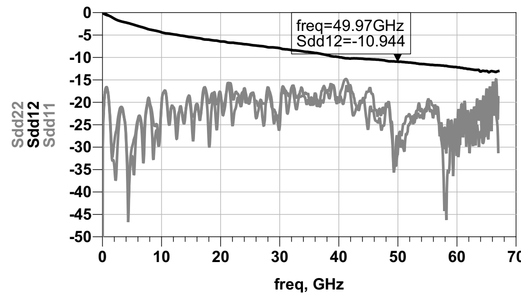


Figure 3.10: Measurement of the coupled tapered trace shown in Figure 3.9, both ends are connected to 1.85 mm *connectors. The center trace pitch is* 150 µm *corresponding to a chip-scale pitch. The total trace length is about* 7.5 cm.

3.3.4 Rat-race coupler

The rat-race coupler shown in Figure 3.11 consists of a 275 µm wide center circle with a radius of 2800 µm, which corresponds to 70.7 Ω on the Rogers RO4003C 221 µm LoPro substrate. The connecting traces are 375 µm wide 50 Ω traces with a Southwest Microwave 2.4 mm connector launch. The measured results are shown in Figure 3.12. It is clear that the rat-race coupler works around 50 GHz. The amplitude difference is below 1 dB and

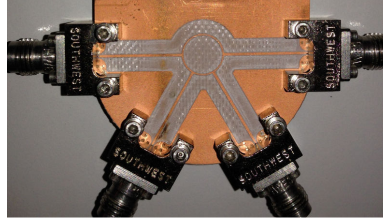


Figure 3.11: Photograph of the measured rat-race coupler using microstrip lines and a coplanar launch.

the phase difference is within 10% of the designed 180° from $47\,\mathrm{GHz}$ to $52\,\mathrm{GHz}$.

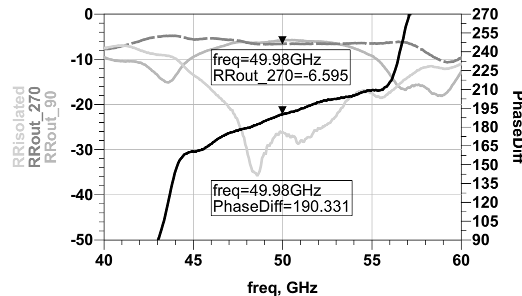


Figure 3.12: Measurement results of the $50\,\mathrm{GHz}$ *rat-race coupler shown in Figure 3.11.*

3.4 Flip chip mounting of the receiver chip

The chips designed in this thesis have a bit rate above 80 Gb/s, which requires each individual component in the channel to exceed a bandwidth of 40 GHz. To accomplish maximum bandwidth in the interconnect, the chips will be directly flip chip mounted on the PCB boards rather than wire bonded to minimize interconnect parasitics. The techniques mainly used for flip chip mounting are reflow soldering, thermocompression bonding and thermosonic bonding.

Reflow soldering

On the chip, NiAu balls are grown which are then soldered to the board. A stencil is made to print solder paste on the board, afterwards the chip is positioned on top of the board and carefully dropped into place. The alignment of the solder mask to the board traces and sten-

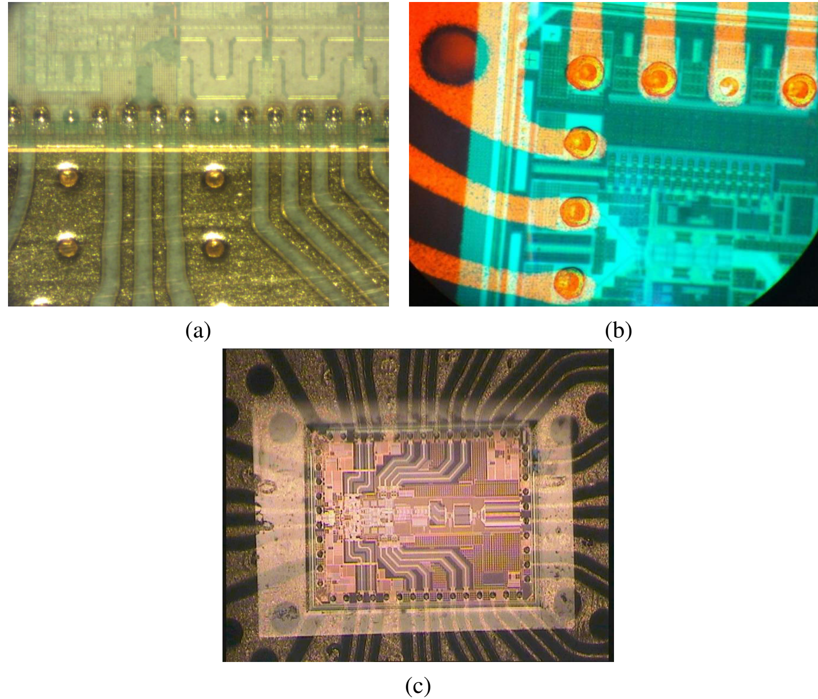


Figure 3.13: Flip techniques (a) Flip chip soldering at cleanroom CMST (b) Thermocompression bonding at imec after stud bumping at Optocap (c) Thermosonic bonding at Optocap

cil positioning is critical. Once the chip is on the board, it is soldered in a reflow oven. The stand-off (height between chip and board) is determined by the amount of solder paste, the solder mask and the height of the electroless NiAu bumps uniformly grown on the chip. This technique was performed by the Centre for Microsystems Technology of Ghent University (CMST) and is shown in Figure 3.13a.

Thermocompression bonding

This mounting method uses soft gold bumps, most of the time generated using a ball bonding machine. The amount of balls on top of each other determines the stand-off. Afterwards, the chip and board are heated, while compressing the chip on the board forming a solid electrical connection. This technique was performed by imec[2] and is shown in Figure 3.13b. For this method no solder mask is needed,

[2] Imec is a micro- and nanoelectronics research center headquartered in Leuven, Belgium. http://www.imec.be

adding solder mask limits the stand-off.

Thermosonic bonding

This technique is comparable to thermocompression bonding but also using ultrasonic energy during the bonding step. Consequently, the compression force and temperature can be reduced, which leads to less thermal and mechanical stress to both substrate and component, reducing the chance of cracking. Consequently, this makes it the preferred method for silicon chips with only bumps on the outside. This technique was performed by Optocap[3] and is shown in Figure 3.13c. For this method no solder mask is needed, adding solder mask limits the stand-off.

In a first effort reflow soldering was chosen, since this technique is most suited for mass production. This mounting method was attempted at CMST. Figure 3.14 shows the NiAu bumps that were grown on the pads of the input buffer test structure. Although multiple attempts were made, with different dies or boards or (amount of) solder paste, the mounted chips could never be tested. There were always some critical connections missing (eg. SPI, bias, input, ...).
These samples were investigated further using X-ray, and optical inspection. The X-ray test at the ”Centre for X-ray Tomography” of Ghent University (UGCT) were inconclusive. As illustrated in Figure 3.15, most of the pads are dark gray, indicating a high concentration of metal. However, this inspection does not guarantee a good connection.
Since this non-destructive method does not show a clear problem, 2D cross sections of one sample were made at CMST as is shown in Figure 3.18. These 2D cross sections were first optically inspected and afterwards a Scanning Electron Microscopy (SEM) diagnosis was performed, shown in Figure 3.16a and 3.16b respectively. Both methods indicated that most NiAu bumps were well soldered to the board however soldermask misalignment caused some bad connections. A closer inspection also showed that the pad aluminum was of limited quality. As is shown in Figure 3.17 the aluminum on the pads showed obvious cracks.

[3]Optocap Ltd provides contract package design and assembly services for microelectronic and optoelectronic devices. http://www.optocap.com/

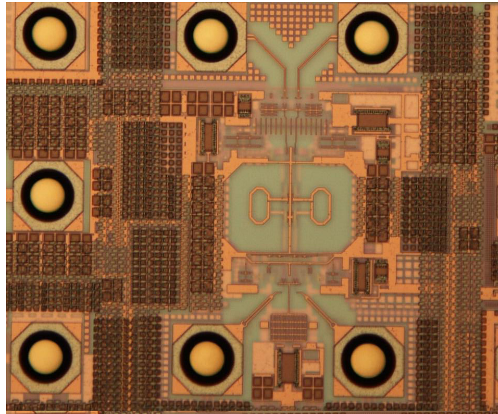


Figure 3.14: Photo of NiAu bumps grown on the chip with test structures.

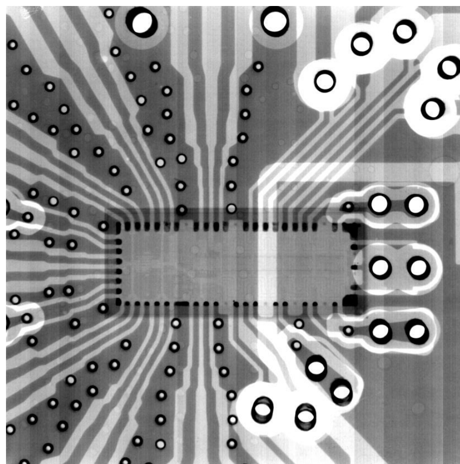
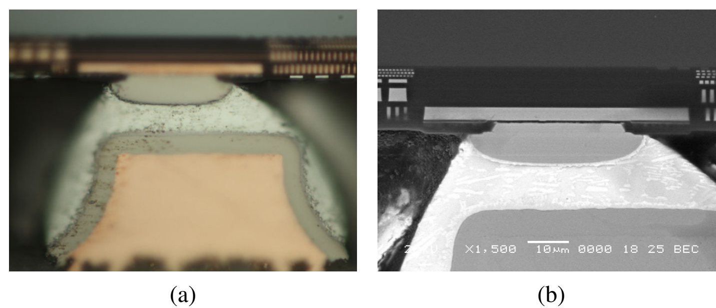


Figure 3.15: Non-destructive check of connectivity by X-ray at UGCT.



(a) (b)

Figure 3.16: Flip chip techniques: (a) optical photo of a bump that should be connected after soldering (taken by CMST), (b) SEM photo of the same bump that should be connected after soldering (taken by CMST).

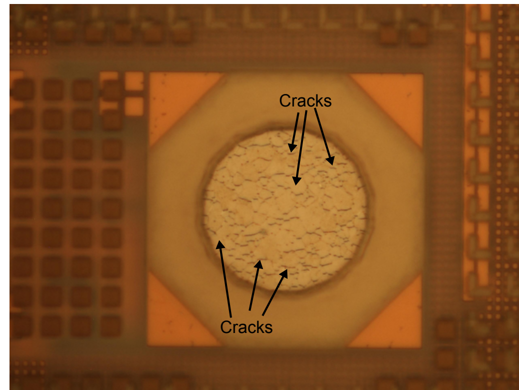


Figure 3.17: Unexpected cracks in the top aluminum layer covering the pads this together with the solder mask misalignment is why soldering was unsuccesful.

After the first unsuccessful trial with reflow soldering, it was opted to perform thermocompression bonding. Imec could do the thermocompression bonding on the NiAu board finish, which was chosen for the chip soldering, but was not able to do the bumping. Optocap was contacted for the mounting, but was not willing to flip on the NiAu board. This resulted in the first samples to be bumped by Optocap and afterwards flipped by imec. This lead to the first real connectorized measurement results.
In the second board run, a soft gold finish was selected and the chips were flipped and bumped by Optocap. It was advised to use thermosonic bonding due to its higher yield when having a limited number of bumps on the die. Both flipping with and without underfill was tried but for mechanical reasons the last samples were all ordered with underfill, since the chips tend to loosen if no underfill is present. Moreover, the electrical benefits of no underfill were limited.

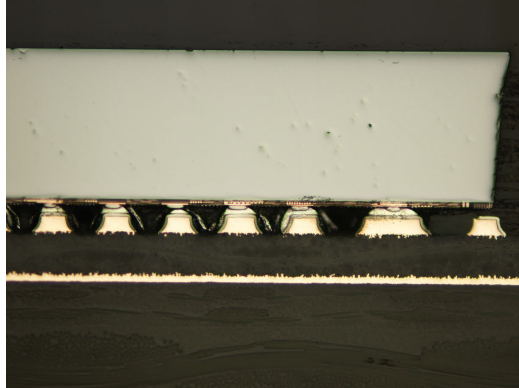


Figure 3.18: Flip chip soldering at cleanroom CMST, cross section.

3.5 Design of a flip chip test board

There were two main revisions of the test board, the first version was designed to test the receiver test structure. The second version was designed to mount the final receiver chip including DEMUX. The second version was a modified board and included changes and optimizations based on the experience obtained from the first board and the test structures. Most of these changes have been discussed in the previous sections, others are related to the chip mounting method and soft gold finish. The biasing was moved to reduce board area and provided by a generic control system designed in the lab and currently used for most chip testing.

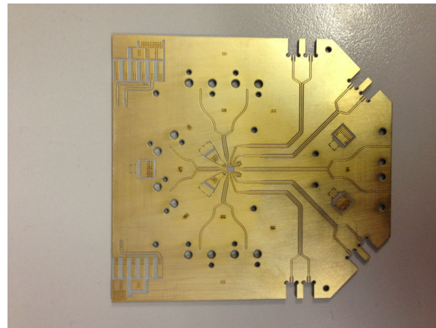


Figure 3.19: Improved test board of OGREv2 using $1.85\,\mathrm{mm}$ *and* $2.4\,\mathrm{mm}$ *connectors.*

An improved board design, making use of $1.85\,\mathrm{mm}$ angle mounted connectors is shown in Figure 3.19. Making use of these connectors, a much

higher bandwidth is achieved. A comparison of the probed results is shown in Figure 3.20. Reflections near the connector have disappeared, resulting in a much smoother transmission characteristic. It should be noted that the used probes had a bandwidth of only 50 GHz. Above this frequency, mode conversions reduce the measurement accuracy.

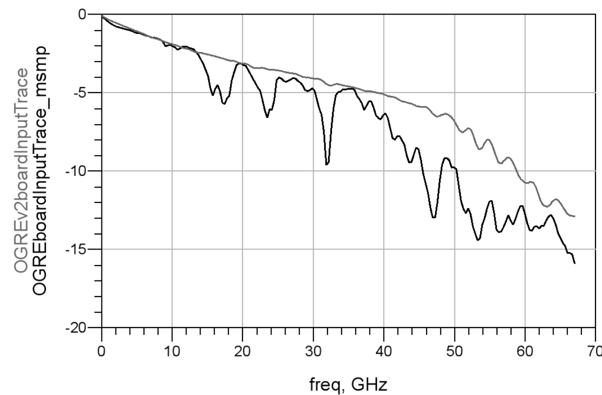


Figure 3.20: Comparison of insertion loss ripple on original and improved test boards, results above 50 GHz are illustrative since the used probes are only mode free up to 50 GHz.

The final test board, shown in Figure 3.19, uses 3 different types of connectors, depending on price, insertion loss, signal bandwidth and board size. For the high-bandwidth in- and outputs the 1.85 mm Rosenberger angle launch connectors are used, in combination with traces that are as short as possible to reduce insertion loss.
As indicated in Section 3.2.3, these connectors are expensive, but have a footprint which allows to put them in the middle of a board. Moreover, they have 67 GHz of bandwidth including footprint. For the clock, 2.4 mm Southwest Microwave end launch connectors are used. As this is a single frequency, bandwidth is of less importance and more frequency dependent loss can be tolerated, which allows the use of longer traces. For the quarter rate DEMUX outputs mini-SMP connectors are used because of their limited size and cost effectiveness. For SPI control signals and biasing, a 2.54 mm pin header is used, surface mount was selected to reduce the amount of different plated through hole (PTH) diameters, which simplifies the plating and by doing so increases the etching control on the 80 μm traces.

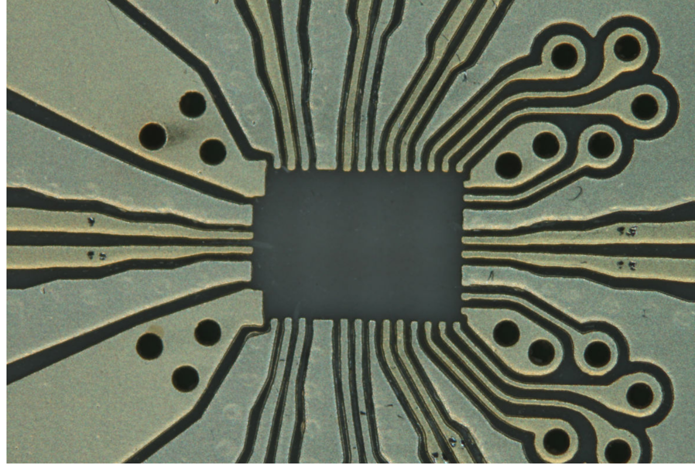


Figure 3.21: Chip footprint in the middle of the receiver board.

The board, shown in Figure 3.22, is built around the footprint of the receiver chip, which is shown in Figure 3.21. Because the chip is flipped directly on top of this board it is very important that the spacing of the pads is equal to the 150 µm pitch on-chip. Depending on the mounting method, other important aspects concerning this footprint are: the thickness of the gold plating, the top width of the lines, and whether or not there is a need for a solder mask. This was discussed in Section 3.4.
Next to these items there were also some alignment markers, in order to do potential adjustment with the milling machine in the lab, and 4 holes for a heat sink were added.

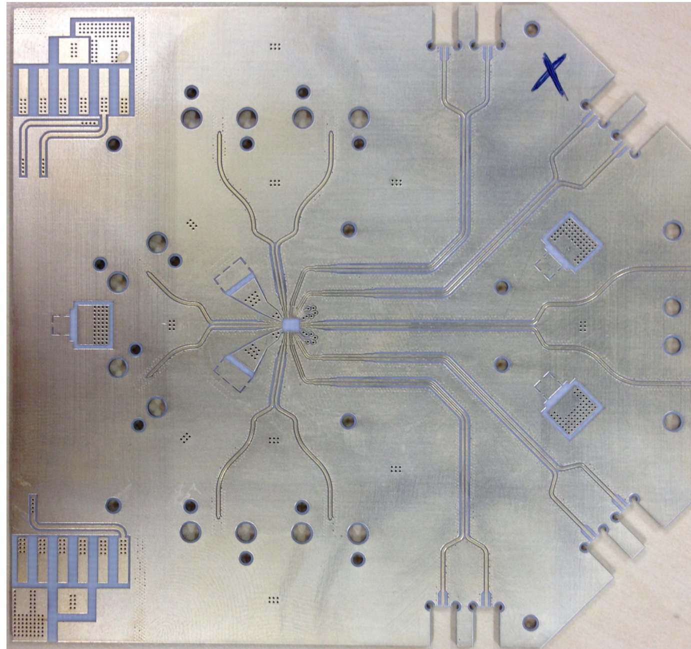


Figure 3.22: Photo of the receiver board without mounted components.

3.6 Design of a multilayer active daughter card

To test more realistic backplane configurations, to allow for crosstalk measurements and to have a system demonstrator, active daughter cards were designed. Each daughter card contains 4 mounted chips, which allow for high-speed backplane link measurements including crosstalk aggressors. A photograph of the finished board is shown in Figure 3.23.

3.6.1 Board stack up

Starting out with the board design, all preconditions were noted down and a board stackup was defined, based on the knowledge available from the design of previous boards and feedback from the board manufacturer. Such preconditions, amongst others, include:

One board

Typical backplane demonstrators have a set of 2 different daughter cards. The cards are typically very similar but with the backplane connector mounted on opposite sides of the board. This is to allow demonstrators having both measurement cables directed toward the

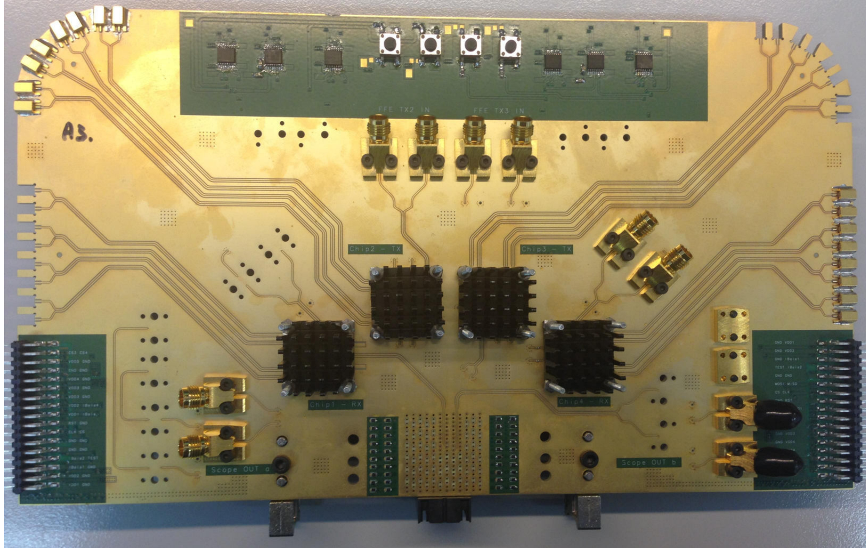


Figure 3.23: Photo of the Designcon daughter card with chips and heatsinks mounted.

outside. This, however, doubles the board manufacturing cost and is prone to errors during design.

This daughter card was designed to fit in both slots. This could be done by using angle mount connectors. In this way measurement cables can connect from the top, left or right of the board. But also careful layout and 3D verification of the routed pairs was done. The routing will be explained in more detail in Section 3.6.2.

Board size

The board needs to contain 4 chips including all control and input output signals. To limit the amount of vias in the high-speed traces, layer transitions should be avoided. Hence, these traces are mainly routed on the top layer (on which the chips are flip chip mounted). To ensure the mechanical stability of a 25 cm by 14.5 cm board, the board thickness should not be less than 3 mm, which is comparable to the passive daughter cards that come with the FCI ExaMAX® backplane demonstrator.

Connector selection

The used connectors will have an influence on the board size and routing topology. The backplane connector is a press fit ExaMAX® connector. All other connectors are surface mount connectors, this to-

gether with the chips being mounted on the top layer, explains why mainly GCPW traces are used on this board.

Routing layers

The amount of routing layers determines the total amount of layers in the board stackup. 12 pairs will be routed from 4 different columns (as is indicated in Figure 3.24). Hence, at least 3 routing layers are required to get all pairs out of the connector. Furthermore, the bottom layer is also required to mount connectors. This results in 2 stripline layers and 2 GCPW layers for high-speed signals. From Figure 3.24, it is also clear that the middle columns will be used for routing the channels connected to the chips.

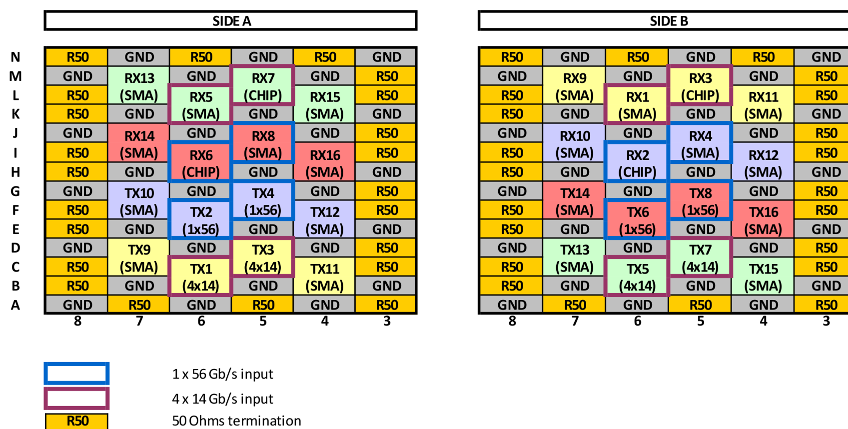


Figure 3.24: Routing example of the ExaMAX® connector.

Material selection and manufacturablity

Low loss material is preferred to increase link length. However, the material should be realistic (with respect to typically used material in these kind of applications) and preferably the same for the entire link. Since the available backplane demonstrator uses Megtron 6 material, the daughter card will also be designed on this material. For symmetry and manufacturability, the same material has to be used as core and prepreg. This is necessary to have similar thermal expansion. In this way, less displacement errors are introduced. The high-speed layers were routed on material cores with a Hyper Very Low Profile (HVLP) copper foil. In between the material cores Megtron 6 prepreg was used.

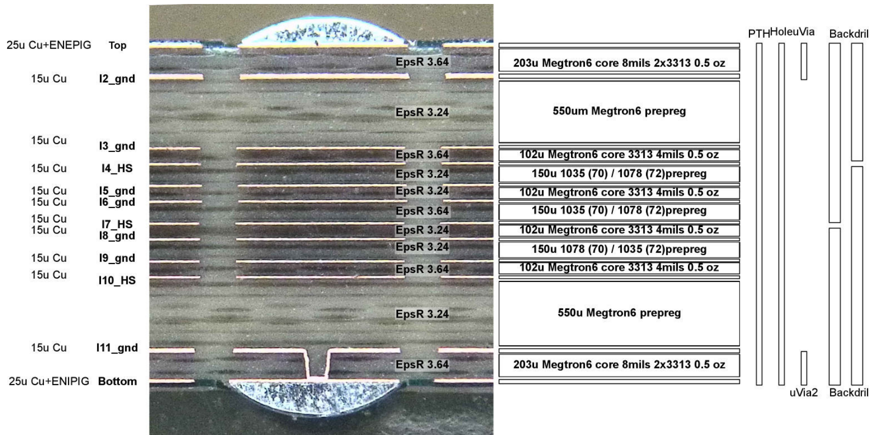


Figure 3.25: Photo of the actual stackup on a 2D cross section.

Taking all these preconditions into account, the stackup, which is shown in Figure 3.25, is obtained. This is a 12 layer stackup using four high-speed routing layers. The top and bottom layers are routed using GCPW lines and layers 4 and 7 use striplines. The stripline substrate has a $102\,\mu\mathrm{m}$ core and a $150\,\mu\mathrm{m}$ prepreg where the top and bottom layers have a $203\,\mu\mathrm{m}$ core. Multiple layers of prepregs, with a combined thickness of $550\,\mu\mathrm{m}$, are used between layers 2-3 and, for symmetry reasons, between 10-11 to increase the total board thickness to about 3 mm. The additional layers are mainly used to create a low impedance ground net but also to transport all kinds of DC control signals and four separate chip supply voltages. To be able to place the microvias as close as possible to the edge of the top layer, they were lasered from layer 2 to 1 and 11 to 12. This results in a higher metal thickness in layers 2 and 11, which are the ground layers. All signal vias

are backdrilled to reduce the stub length and, hence, maximize the trace bandwidth.

3.6.2 Routing

There are 4 chips mounted on the board, 2 transmitter chips and 2 receiver chips. Each transmitter chip (described in more detail in the PhD dissertation of my colleague Yu Ban) has 2 outputs that drive a backplane channel, one connected to a 4-to-1 MUX and a feed forward equalizer (FFE) and an other one just connected to the output of an FFE. These 2 chips will drive 4 pairs in the ExaMAX® connector.

Additionally, 2 receiver chips are mounted on the board, each connected to a transmitter chip through the channel (with a broadband capacitor soldered in between). One receiver is connected to a FFE-only output of one of the transmitter chips, the other receiver chip is connected (through the channel) to a 4-to-1 MUX with FFE output.

ExaMAX® routing

All routed signal pairs of the ExaMAX® connector, illustrated in figure should have the highest possible bandwidth, therefore they are carefully laid out. In total 12 pairs are routed, four of which are connected to the transmitter chips. Two are connected to the receiver chips. And the other 6 are connected to a set of $1.85\,\mathrm{mm}$ Rosenberger connectors. This can be seen in Figure 3.24.Connecting some of the pairs directly to $1.85\,\mathrm{mm}$ connectors allows easy debugging of the channel in the frequency domain and inspection of the eye in the time domain. It also allows the addition of extra cross talk channels.

The pairs that are not routed but which are next to routed pairs are terminated with $50\,\Omega$ resistors to reduce reflections. In Figure 3.26, the backplane routing of one of the middle columns of the ExaMAX® connector is schematically represented.

Full rate signal routing

All full rate signals coming from or going towards the chip are routed using as little layer transitions possible. On the top and bottom layer, GCPW are used to connect to 1.85mm connectors. On the middle layers, mainly used for connecting to the ExaMAX® connector, differential stripline traces are used. The vias and launches on these traces were optimized using CST Microwave Studio with the model shown in Figure 3.27, based on the methodology previously explained in Section 3.2.3. Next to the high-speed data traces, a half

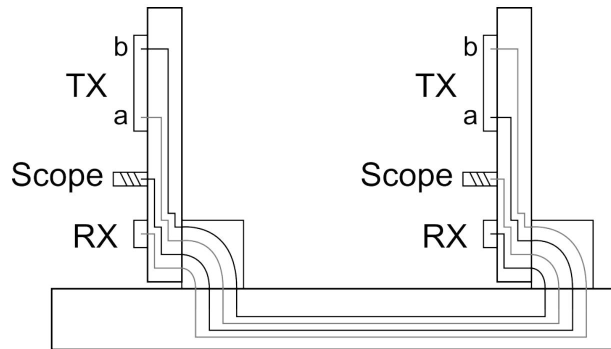


Figure 3.26: Routing example of one of the middle columns, not to scale.

rate clock has to be provided to each chip, for these traces the most important specification is equal length to keep the 180° phase relationship of the differential clock.

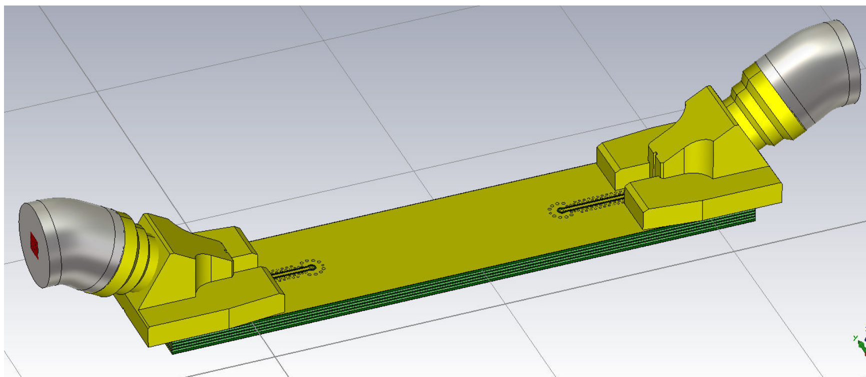


Figure 3.27: Simulation of via performance using realistic launches.

Quarter rate routing

Next to the full rate in- and outputs of the chips, each chip also has four quarter rate differential signals, making a total of 16 differential traces which are routed to miniSMP pairs. These are retimed signals with low jitter and high-amplitude and run at quarter rate so the routing is less critical. They are routed in such a way that the amount of board real estate is minimized.

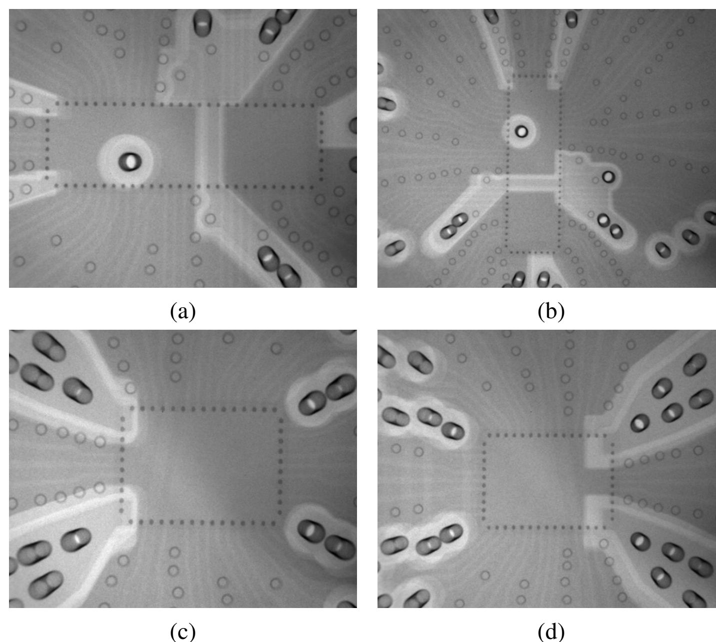
Power and control signals

The routing of the control signals and the power and ground connections was mainly done by Joris Van Kerrebrouck. He also made the

necessary adjustments to the test platform software and hardware to be able to control the 4 chips mounted on one daughter card.

3.6.3 Mounting order

After board production, electrical verification and optical inspection of the critical features is done in the lab before the chips are mounted. Next, the chips are mounted on the board and the connections are verified using X-ray inspection. The resulting pictures are shown in Figure 3.28.

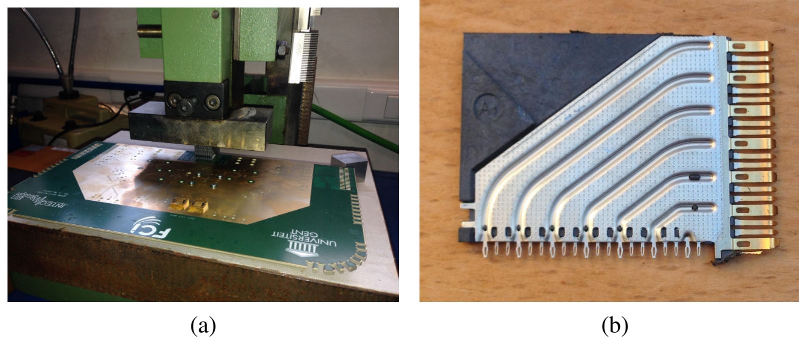


(a) (b) (c) (d)

Figure 3.28: X-ray inspection by Optocap of the thermocompression mounting of the chips on the board.

After electrical verification and mounting of the heat sinks, the backplane connector is pressed onto the board, as is shown in Figure 3.29a. The ExaMAX® connector contains multiple A and B columns next to each other, called IMLAs, one IMLA (type A) is shown in Figure 3.29b. It is clear that the ground pins are larger than the signal pins. Using smaller signal pins allows for smaller signal vias and therefore better high-speed performance.

When all chips and the ExaMAX® connector are successfully mounted, the



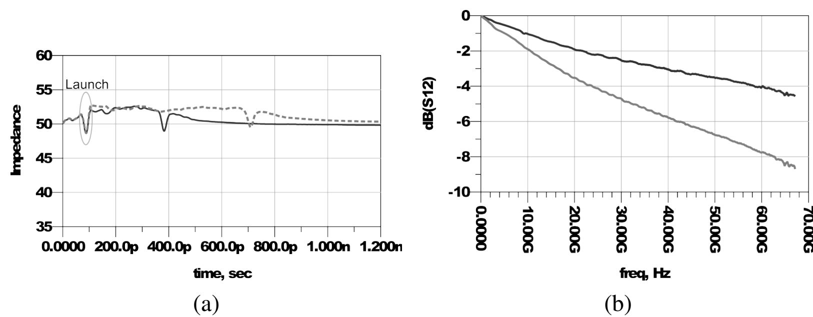
(a) (b)

Figure 3.29: Press used to mount ExaMAX® connector (a) and IMLA of an ExaMAX® connector (b)

remaining connectors and control signal buffers are soldered or screwed on and the boards can be tested.

3.6.4 Board performance

On the board, several test traces are available. To optimize board space usage, the footprint of the $1.85\,\mathrm{mm}$ angle mount connector was reused but with the connector mounted on the back (or 180° rotated).

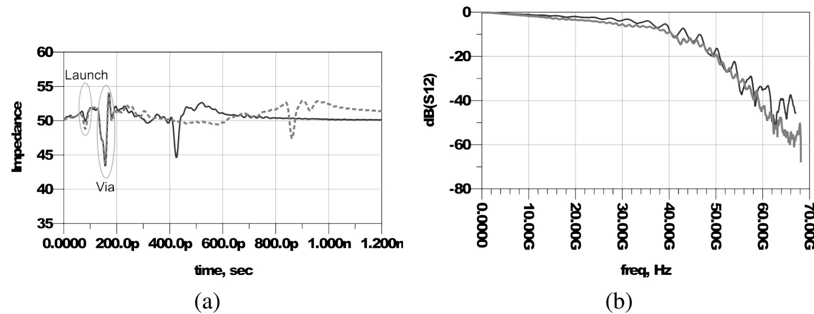


(a) (b)

Figure 3.30: TDR (a) and insertion loss (IL) (b) of two traces with different length routed on the top layer.

In Figure 3.30, two different top layer trace lengths are compared. From these measurements the insertion loss (IL) of the traces on the top layer is estimated to be 0.65dB/cm (1.65dB/inch) at 28 GHz, and the insertion loss of the connector launch is estimated to be 0.23dB (also at 28 GHz). The IL graph shows a very smooth response indicating little reflections, which is

confirmed by the TDR in Figure 3.30a, which shows good footprint matching.



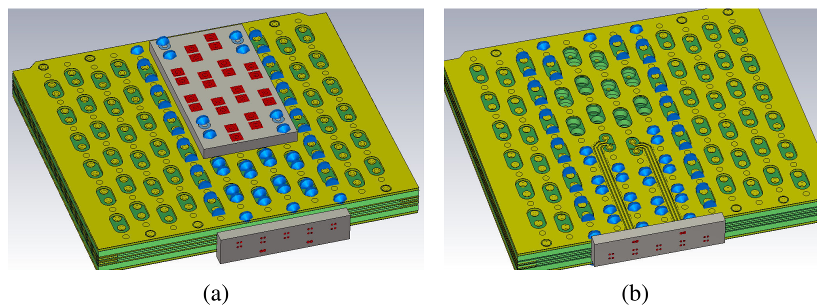
(a) (b)

Figure 3.31: TDR (a) and IL (b) of two traces on one of the middle high-speed routing layers.

The stripline trace perfomance on the middle layers is compared in Figure 3.31. From the IL graph it is clear that the bandwidth is limited to around 40 GHz. The TDR in Figure 3.31a shows that this is due to reflections at the via. The loss of the traces on the middle layers is estimated to be 0.67 dB/cm (1.7 dB/inch) at 28 GHz.

Next to these test structures, two equal lines but with a different amount of vias were available on the board. These allow to calculate the loss of a top to bottom via which is 0.5 dB at 28 GHz.

The performance of the footprint of the ExaMAX® connector was simulated in CST by Jan De Geest, a model can be seen in Figure 3.32. Measurements done using these boards are shown in Chapter 5.



(a) (b)

Figure 3.32: CST Microwave Studio model of the ExaMAX® footprint simulated by Jan De Geest, top view (a) and bottom view (b).

3.7 Conclusion

In this chapter, broadband measurements up to 67 GHz of different high-bandwidth test structures were presented. These structures show that mounted, in contrast to traditional probe, testing of millimeter wave integrated circuits is possible, but also show the possibility to use these devices, mounted on a board for system experiments in a more complex setup. The difference in loss between the Rogers RO4003C LoPro and NoLoPro materials, 1.1 dB and 1.4 dB per centimeter at 50 GHz respectively, was shown, as well as a comparison between screw-on connectors from Rosenberger and Southwest Microwave. Coupled traces, tapered down to 150 µm pitch, allowing to directly flip a chip on the traces and a 50 GHz rat-race coupler to convert a single-ended clock into a differential clock, prove the feasibility of a high-speed, board mounted measurement setup.
At 50 GHz and above, LoPro material has a clear advantage over NoLoPro material because the loss per centimeter is about 25% lower. At 30 GHz and below, however, there was hardly any difference in loss between the two materials. Comparing the 1.85 mm Rosenberger connector to the 2.4 mm Southwest Microwave connector showed that the 2.4 mm connector has about two times the insertion loss at 60% of their maximum frequency.
For research projects the thermosonic bonding flip chip approach seems the most reliable, it can be preformed at Optocap if the PCB has a soft gold finish and correct footprint. For this technique there is also the additional benefit that no solder mask is required around the chip footprint.

References

[1] T. Demeester, D. Vande Ginste, and D. De Zutter. Accurate study of the electromagnetic and circuit behavior of finite conducting wedges and interconnects with arbitrary cross-sections. In *2010 IEEE 19th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS)*, pages 133–136. IEEE, 2010.

[2] B. Rosas. Optimizing test boards for 50 GHz end launch connectors: Grounded coplanar launches and through lines on 30 mil rogers ro4350b with comparison to microstrip. *Southwest Microwave, Inc., Tempe, AZ*, 2007.

[3] Recommended pcb layout 1.85 mm connector (mb 389). *Rosenberger Corp.*, 2012.

[4] Southwest Microwave Inc, 2.40mm jack (female) end launch connector (1492-03A-5), 2012

[5] K. Chang and L.H. Hsieh. *Microwave ring circuits and related structures*, volume 156. John Wiley & Sons, 2004.

Part II

Measurements and results

4

84 Gb/s SiGe BiCMOS duobinary serial data link including Serialiser/Deserialiser (SERDES) and 5-tap FFE

**Timothy De Keulenaer, Guy Torfs, Yu Ban,
Ramses Pierco, Renato Vaernewyck, Arno Vyncke,
Zhisheng Li, Jeffrey H. Sinsky, Bartek Kozicki
Xin Yin and Johan Bauwelinck**

Based on the article published in Electronics Letters

This work was supported by The Agency for Innovation by Science and Technology in Flanders (IWT).

⋆ ⋆ ⋆

THE increasing demand for bandwidth fuels the development towards high data rate electrical serial links. These links generally suffer from

considerable frequency-dependent loss, introducing the need for equalization at 10 Gb/s and higher. Modulation schemes with improved spectral efficiency, with respect to non-return to zero (NRZ), combined with feed-forward equalization (FFE), allow increasing the chip-to-chip data rate with the drawback of a more complex, e.g. multi-level, receiver (Rx). The use of duobinary modulation (DB) is presented to realize a high-speed serial link. The increase in complexity of a DB Rx is limited, whereas the required channel bandwidth compared with NRZ is reduced. Furthermore, the need for equalization when compared with PAM4 is reduced as the required roll-off that is needed to create a duobinary modulated signal from an NRZ stream can incorporate the frequency-dependent loss of the link.

4.1 Introduction

In this paper, a transmitter (Tx) and receiver (Rx) chipset targeting a serial data rate of 84 Gb/s is presented. A block diagram of the complete transceiver chain is shown in Figure 4.1. The Tx consists of a 4:1 multiplexer (MUX) connected to a 5 tap FFE with an approximate delay of 12.5 ps between the taps. Together with the channel, the FFE creates an equivalent channel that transforms the NRZ from the MUX into a duobinary modulated signal at the Rx input (Figure 4.1). The receiver chip includes a duobinary front end connected to a 1:4 demultiplexer (DEMUX). A data rate of 84 Gb/s is achieved across a differential link which includes a parallel pair of 10 cm coax cable and the 5 cm grounded coplanar waveguide (GCPW) traces on the chip test boards. We believe this to be the fastest reported electrical duobinary modulated transmission experiment to date [1]. The Tx and Rx chips are fabricated in 130 nm STMicroelectronics SiGe BiCMOS 9MW technology and mounted directly onto the PCB substrate using thermosonic flip-chip bonding. SPI is used to control both chips and to adjust the bias currents, the comparison levels of the receiver, the tap weights of the FFE and the clock delays in the MUX and DEMUX.

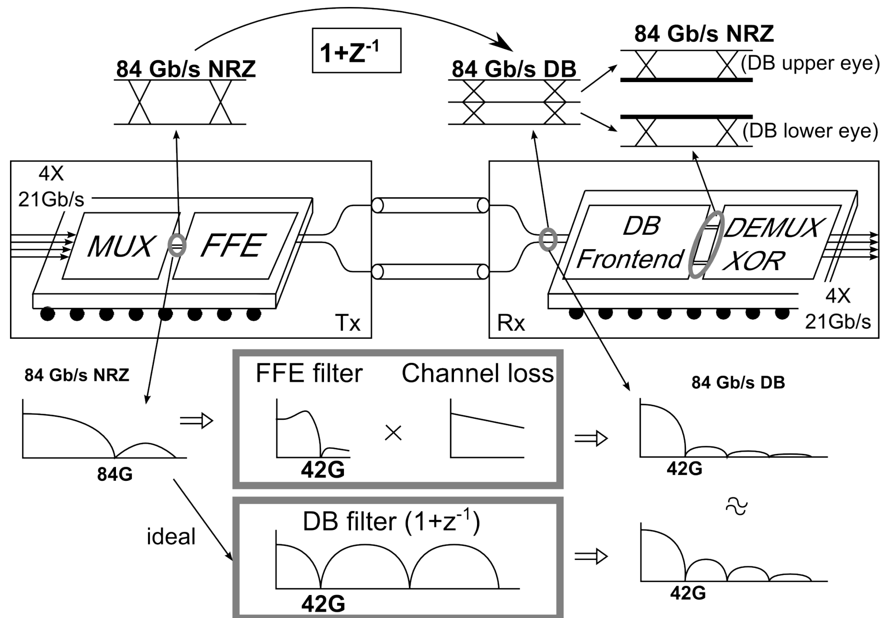


Figure 4.1: Architecture of the presented transceiver chain (middle), the time domain transition from NRZ to duobinary modulation (top) and the spectral shaping of the channel to receive duobinary modulation (bottom).

4.2 Transmitter

The 4:1 MUX has a tree architecture with the final 2:1 selector stage feeding both the FFE driver and the MUX test output driver. To overcome problems related to limited setup and hold time of the input retimer and to facilitate data alignment, the MUX input data can be delayed on-chip by 12.5 ps and the clock phase of the half-rate multiplexers can be selected on-chip. The MUX provides a clean NRZ waveform at 84 Gb/s as shown in Figure 4.6 by the eye diagram measured at the MUX output of the Tx test board.
The 5 tap FFE topology is shown in Figure 4.2. The delays are implemented using meandered on-chip transmission lines with a length of 750 um in between the gain cells. The measured delay between the cells, including line loading, is 12.4 ps ($\pm$0.1 ps). The gain cells (bottom Figure 4.2), are based on a Gilbert cell architecture of which the gain is tuned linearly by changing the tail current difference between both differential pairs using an 8-bit monotonic DAC. By keeping the summed current of both differential pairs constant, the current flowing through the transmission line termination resistors is constant, which in turn keeps the bias voltage of the FFE output

buffer constant.

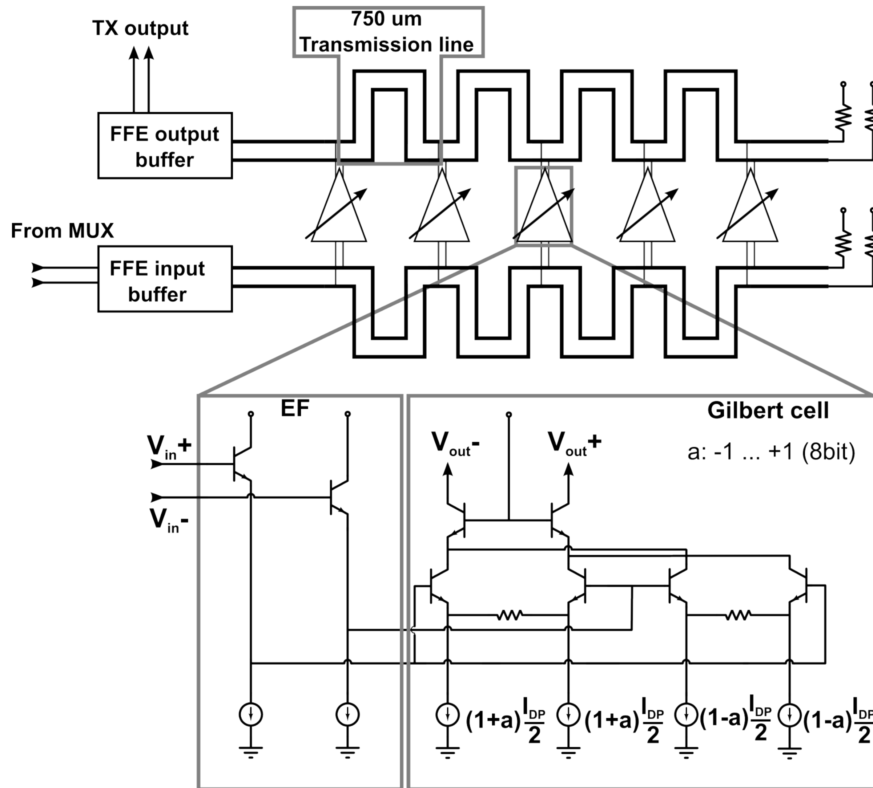


Figure 4.2: Overview of the 5 tap FFE (top) and the circuit of the gain cell (bottom).

The Gilbert cell layout is relatively small compared to the meandered transmission lines. By splitting the gain cell, placing the emitter followers as an input buffer at the input transmission line and placing the actual Gilbert cell at the output transmission line, long leads and excessive loading of the transmission lines are avoided.

4.3 Receiver

The receiver block diagram is given in Figure 4.3. A transimpedance amplifier (TIA), with inductive peaking to increase the bandwidth, is used as a matched linear input buffer which allows achieving both noise and impedance matching simultaneously [2]. The output of the TIA is connected to two parallel level-shifting limiting amplifiers (LSLA), which con-

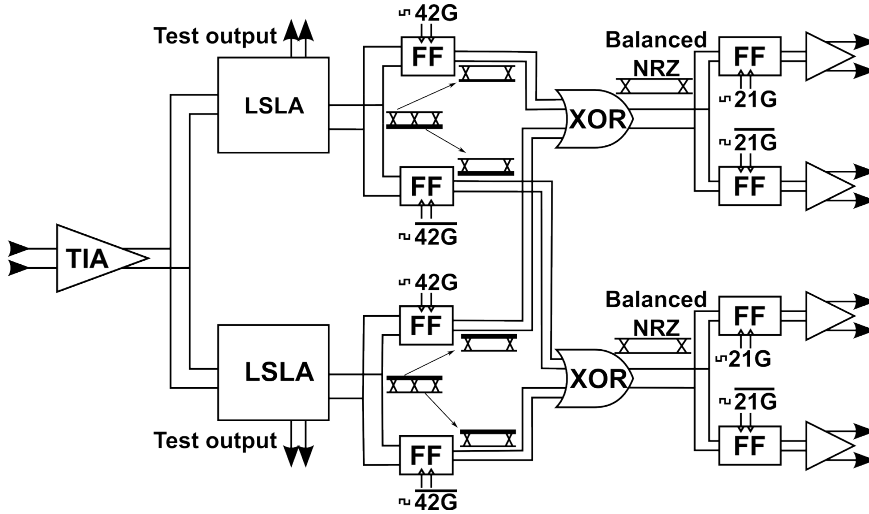


Figure 4.3: Block diagram of the Rx chip and an illustration of how the unbalanced NRZ streams are first demultiplexed before they are decoded by the XOR gate.

vert the DB stream into two NRZ-like streams, recovering the upper and lower eye of the DB signal. A block diagram of the LSLA is shown in Figure 4.4. The level comparison is done in 2 stages having 8 and 4 bits of digital control respectively. This allows re-enforcing the comparison level through the signal path. This is needed because the comparison of one of the duobinary eyes leads to mark densities that deviate from 50% (the theoretical densities are 25% and 75% with balanced input data). The comparison level is set by tuning part of the emitter follower (EF) current through the resistors R_E. Shunt capacitors C_E are added to short R_E in the data path (Figure 4.4). An asynchronous XOR introduces considerable jitter, hence the DEMUX first clocks the two LSLA outputs to four data streams with half the bit rate. At this lower bit rate, the XOR operation is performed to decode the DB stream into two balanced NRZ streams (Figure 4.3), followed by an additional demultiplexing operation resulting in four quarter rate NRZ data streams.

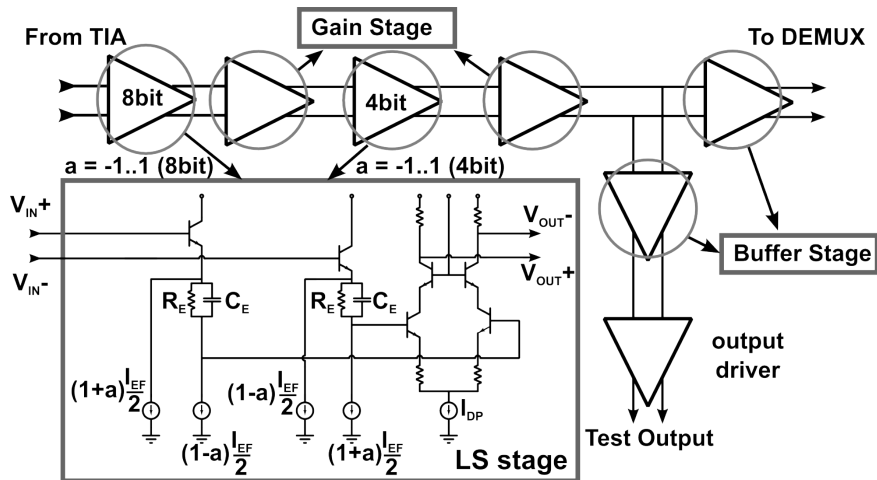


Figure 4.4: Block diagram of the 5-stage LSLA (top) and schematic of the level-shifting stages (bottom left).

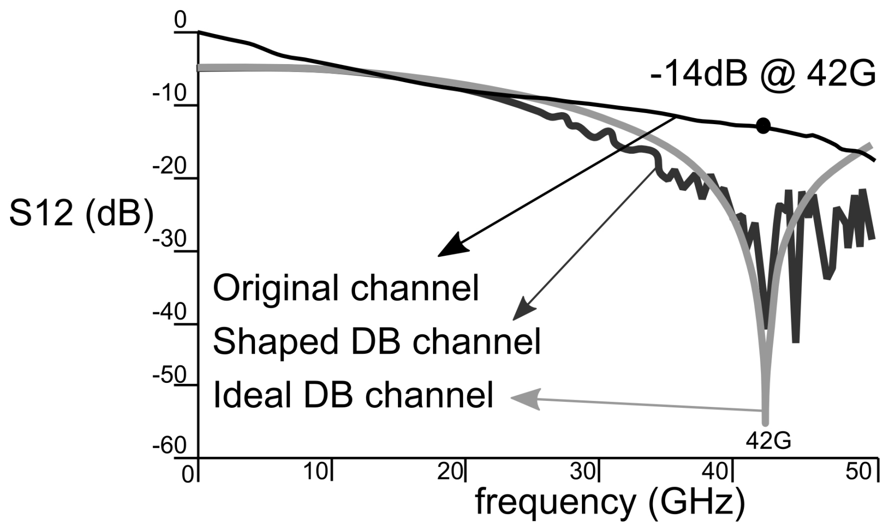


Figure 4.5: Total channel loss (black), the ideal duobinary shape (grey) and the duobinary shaped channel generated by optimizing the tap weights of the FFE (dark grey).

4.4 Measurements and conclusion

The chips are mounted directly onto a Rogers RO4003C substrate using thermosonic flip-chip bonding. The traces, together with 1.85 mm screw-on angle mount connectors, were designed to have a smooth frequency response up to 67 GHz and above [3]. The Tx and Rx test boards are interconnected with two phase-matched 10 cm coaxial cables and 1.85 mm DC blocks. The total combined channel loss at the Nyquist frequency (42 GHz) is 14 dB as can be seen in Figure 4.5. Four identical PRBS7 streams with 0, 63, 95 and 31 bit delays respectively are combined, resulting in a full-rate PRBS at the output of the MUX. This prevents the need for precoding since a precoded PRBS yields the same PRBS, simply shifted in time. The full-rate DB eye at the output of the 10 cm coax cable is shown in Figure 4.6 together with the NRZ eye at the output of the MUX.

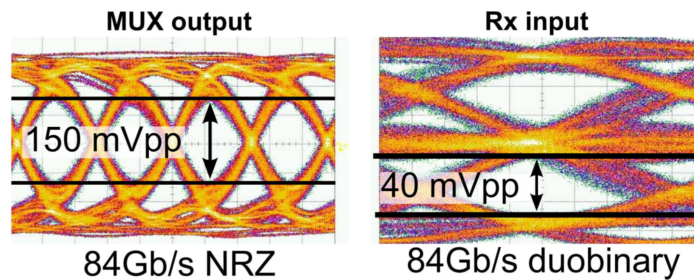


Figure 4.6: 84 Gb/s output eye diagram of the MUX (left) and the 84 Gb/s duobinary eye at the input of the receiver.

The BER measured on a DEMUX output channel showed successful data transmission of 84 Gb/s across a serial link with a BER of 5.3×10^{-12}. The Rx chip with DEMUX and the Tx chip including MUX consume 1.25 W and 750 mW respectively from a 2.5 V supply and occupy 1.55x4.59 mm^2 and 1.93x2.58 mm^2 respectively. This results in an overall power consumption of 24 mW/Gb/s for the 84 Gb/s duobinary modulated link, showing that the presented Rx and Tx operate at very high data rate with reasonable power consumption. Chips with similar complexity, but working at lower data rates, are presented in [4] and [5]. In [4] an efficiency of 43 mW/Gb/s is shown for a 44 Gb/s link including CDR and SFI-5.2 interface, while [5] showed a low power consumption of 11 mW/Gb/s at 40 Gb/s and including 5 tap FFE and 3 tap adaptive FIR filter as a front end equalizer, however, without MUX and at much lower rate. To conclude, our work shows that duobinary modulation is a viable alternative for NRZ and PAM4 and that generating and decoding duobinary signaling can be done with reasonable

circuit complexity and power consumption at bit rates above 80 Gb/s.

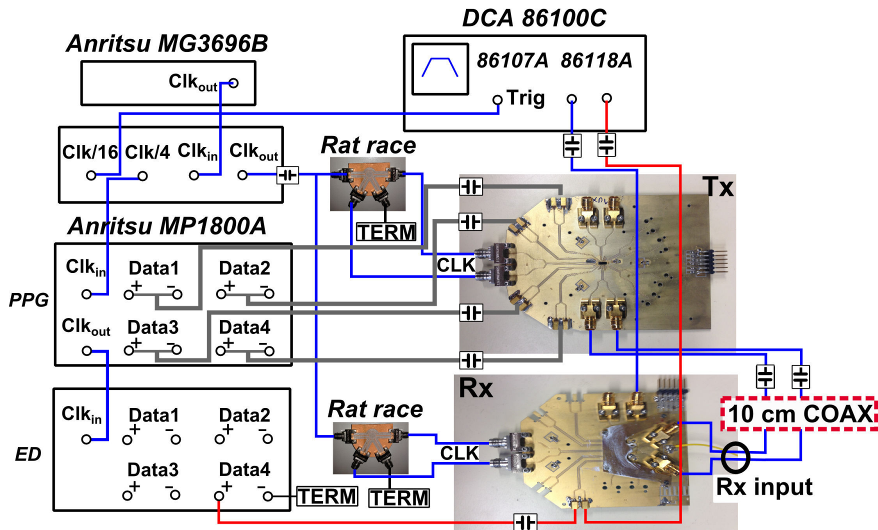


Figure 4.7: Overview of the 84 Gb/s test setup.

References

[1] J.H. Sinsky, A. Konczykowska, A. Adamiecki, F. Jorge, and M. Duelk. 39.4 Gb/s duobinary transmission over 24.4 m of coaxial cable using a custom indium phosphide duobinary-to-binary converter integrated circuit. *IEEE Transactions on Microwave Theory and Techniques*, 56(12):3162–3169, 2008.

[2] S. Shahramian, C. Carusone, and S. Voinigescu. Design methodology for a 40 GSamples/s track and hold amplifier in 0.18μm SiGe bicmos technology. *IEEE Journal of Solid-State Circuits*, 41(10):2233–2240, 2006.

[3] T. De Keulenaer, Y. Ban, G. Torfs, S. Sercu, J. De Geest, and J. Bauwelinck. Measurements of millimeter wave test structures for high speed chip testing. In *2014 IEEE 18th Workshop on Signal and Power Integrity (SPI)*, pages 1–4. IEEE, 2014.

[4] B. Raghavan, D. Cui, U. Singh, H. Maarefi, D. Pi, A. Vasani, Z.C. Huang, B. Catli, A. Momtaz, and J. Cao. A sub-2 W 39.8–44.6 Gb/s transmitter and receiver chipset with sfi-5.2 interface in 40nm cmos. *IEEE Journal of Solid-State Circuits*, 48(12):3219–3228, 2013.

[5] MS. Chen, Y.N. Shih, CL. Lin, H.W. Hung, and J. Lee. A 40 Gb/s tx and rx chip set in 65nm cmos. In *2011 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 146–148. IEEE, 2011.

5

56+ Gb/s serial transmission using duobinary signaling

Timothy De Keulenaer, Jan De Geest, Guy Torfs, Johan Bauwelinck, Yu Ban, Jeffrey H. Sinsky and Bartek Kozicki

Based on the article presented at DesignCon 2015

The work described in this paper has been sponsored by the Flemish agency for Innovation by Science and Technology (IWT). The authors would also like to acknowledge Integrand Software for access to their EM solver EMX.

⋆ ⋆ ⋆

IN this paper we present duobinary signaling as an alternative for signaling schemes like PAM4 and Ensemble NRZ that are currently being considered as ways to achieve data rates of 56 Gb/s over copper. At the system level, the design includes a custom transceiver ASIC. The transmitter is capable of equalizing 56 Gb/s non-return to zero (NRZ) signals into a

duobinary modulated response at the output of the channel. The receiver includes dedicated hardware to decode the duobinary signal. This transceiver is used to demonstrate error-free transmission over different PCB channel lengths up to 50 cm including a state-of-the-art Megtron 6 backplane demonstrator.

5.1 Introduction

Recently, standards groups like the OIF CEI-56G-VSR/MR and the IEEE P802.3bs 400 GbE have been looking into serial data rates above 50 Gb/s as the line rate of future generation PHYs. The OIF is looking at serial data rates of 56 Gb/s, and different signaling and modulation schemes such as PAM4 and Ensemble NRZ are being considered as ways to achieve these data rates over copper.

In this paper we present duobinary signaling as an alternative for PAM4 or Ensemble NRZ. Duobinary signaling is a 3-level modulation scheme that halves the required bandwidth of the channel compared to NRZ and as such has the same bandwidth requirement as PAM4. The generation of a duobinary signal can make use of the inherent frequency dependent channel loss, and hence requires less equalization, greatly reducing the overall system requirements. We will look at the differences between duobinary modulation and PAM4 modulation with respect to transmitter and receiver complexity, required equalization, power consumption etc.

A duobinary transmitter and receiver capable of operating at speeds of 56 Gb/s and above were designed for system-level demonstration of backplane transmission. The transmitter accepts a pre-coded NRZ signal and equalizes the frequency dependent channel loss to produce a duobinary modulated signal at the output of the channel, i.e. the input of the receiver. This dedicated receiver recovers the two eye-patterns typical for a duobinary constellation and decodes them to the original NRZ data sequence.

For system-level evaluation and validation, test boards were designed on which the integrated circuits are mounted using a flip-chip process. Eye-pattern and bit-error-rate (BER) measurements were performed at 56 Gb/s on a state-of-the-art Megtron 6 backplane demonstrator.

5.2 Duobinary signaling

In the quest to reach higher serial data rates in electrical interconnects, there are several active solutions currently being pursued by the industry: increasing serial symbol rates with adaptive equalization and pre-emphasis, shifting to more complex signal constellations [1, 2], and using multi-carrier modulation [3, 4].
Here, as a point of reference for discussion, we use standard 2-level non-return to zero (NRZ) modulation, illustrated in the left-hand side of Figure 5.1. The use of a higher order modulation, such as pulse-amplitude modulation (PAM) with 4 levels (PAM4 illustrated in the center of Figure 5.1), while targeting the same data rate as the NRZ signal, reduces the transmission bandwidth from 1/T to 1/2T. PAM with 4 or more levels (PAM4, PAM5, PAM8, etc.) has been investigated.
The consequence of moving to multilevel modulation is that signal reception requires more decision levels with reduced level spacing. As a result, for the same average signal power and receiver noise, the error probability when receiving a symbol is higher and the signal is more susceptible to deterioration due to inter-symbol interference (ISI). This signal-to-noise ratio (SNR) performance conclusion is based on a channel that has a flat frequency response over the bandwidth of all signaling types being compared. If this was the case, there would be no motivation to use a higher order constellation. In fact, for backplane channels, the amplitude response inevitably rolls off as a function of frequency and will have nulls originating from e.g. via holes between signal layers. As a result, for certain channels, it is possible for multi-level, narrow bandwidth signaling to obtain a larger eye opening, and hence a better SNR than NRZ. Whether or not this happens and for what type of signaling, very much depends on the channel frequency response and the desired data transmission rate. This phenomenon is true for PAM signaling as well as partial response signaling.
PAM4 with equalization is currently being used in the industry and has been shown to provide very good performance even over long traces. However, the susceptibility of PAM4 to ISI typically results in complex and difficult-to-integrate transceiver circuits. The second solution towards reaching higher interconnect speeds is increasing the symbol rate, as indicated on the right-hand side of Figure 5.1. The reduced symbol time (T/2 instead of T) leads to a bandwidth expansion (see Figure 5.2, left and center). However, low cost dielectric materials used to construct backplane printed circuit boards exhibit strong frequency-dependent losses. The frequency dependence leads to deterioration of signal integrity of any signal propagating in that channel. In order to overcome this problem in high-speed interconnect systems,

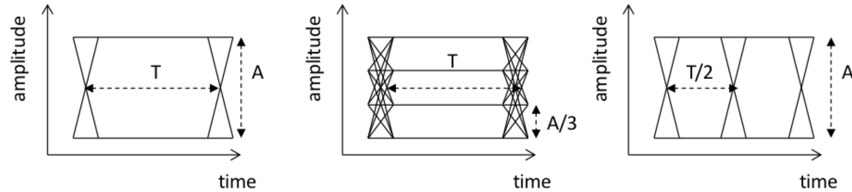


Figure 5.1: Ideal waveforms of three modulation formats (from left to right): standard NRZ, PAM4 and NRZ with double bit rate.

the use of costly microwave substrates and special high-bandwidth backplane connectors is often required. Nevertheless, for long trace lengths, impedance discontinuities from structures like via holes may still result in unacceptable transmission characteristics [5, 6]. In order to overcome the imperfection of the channel, typically equalization and pre-emphasis techniques must correct the entire frequency spectrum of the NRZ data.
The third alternative is to move away from baseband modulation formats towards multi carrier formats. As illustrated in the right-hand side of Figure 5.2, data is transmitted in multiple signal bands, which are individually equalized to accommodate imperfections of the physical channel. Multi carrier techniques have been shown to be practical solutions for last mile digital subscriber loop (DSL) solutions [3]. However, for very high-speed backplane transmission, the cost and power consumption of transceiver integrated circuits still exceed the respective budgets [4]. Multilane transmission techniques for reaching higher interconnect speeds have also been demonstrated [7] but it remains to be seen whether such designs can find a practical application.

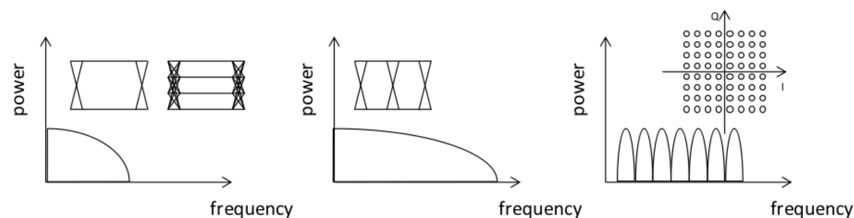


Figure 5.2: Waveforms of three modulation formats: standard NRZ and PAM4 (left), NRZ with higher bit rate (center), multicarrier signal spectrum with quadrature amplitude modulation (QAM)-64 constellation in the inset (right).

The alternative approach to obtaining a higher interconnect speed adopted in this design is the use of partial response signaling. In partial response

formats, the data to be transmitted is temporally distributed over multiple symbols. In the particular case of duobinary modulation, each bit of information is distributed between two symbols as can be expressed by the simple Z-transform filter representation, $1 + z^{-1}$. The controlled ISI forms a 3-level signal. A typical waveform of a duobinary modulated signal is schematically depicted in the center of Figure 5.3.

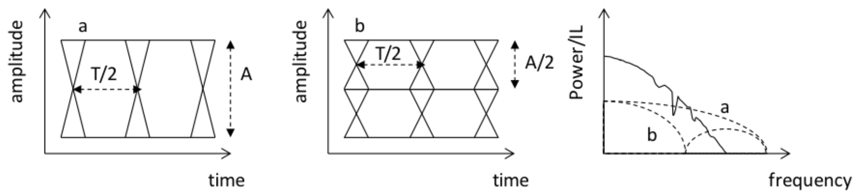


Figure 5.3: Waveform of duobinary modulation (center) and NRZ (left) for the same bit rate. Right hand side of the figure compares the power spectral density (PSD) of NRZ (a) and duobinary (b) modulation for the same rate with a superimposed example insertion loss profile of a physical channel (solid line).

Duobinary signaling was first proposed by Lender in 1963 [8] and evolved over the following decades [9, 10]. This partial response coding technique reduces the required signal bandwidth for transmission compared to NRZ signaling. As a result, the power spectral density (PSD) of the signal is concentrated in the lower frequency region of the channel, which exhibits less loss and irregularities. The cumulative PSD for duobinary and PAM4 is compared to that of NRZ in Figure 5.3. The narrow bandwidth characteristic of the duobinary modulation has been used in both electrical [11, 12] and optical transmission systems [13–17].
Traditionally, binary data is converted to duobinary data at the transmitter and then sent through the channel. In such a system, the conversion to duobinary is done using either a finite impulse response (FIR) filter that takes the form of a delay-and-add filter or a low-pass filter that results in an approximation of this frequency response. The resulting duobinary waveform uses significantly less bandwidth than its binary counterpart. This is clearly seen in Figure 5.4, where the PSD of NRZ and duobinary are compared. It should be noted that in order to increase the signaling rate for duobinary as well as for NRZ modulation, it is necessary to use a higher symbol rate (see Figure 5.3, left and center). However, despite the higher symbol rate the PSD of duobinary remains confined, because not all signal-state transitions occur. By design a duobinary signal will never have a transition immediately from the top to the bottom level.
The required duobinary filter response can also be realized using the combi-

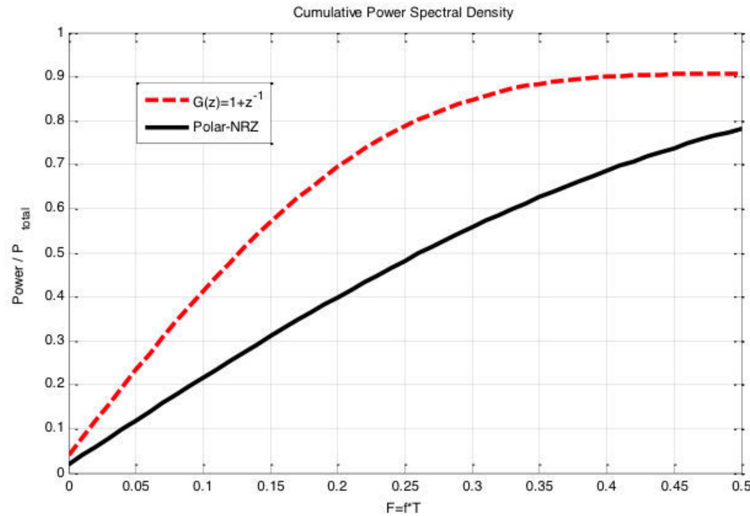


Figure 5.4: Cumulative PSD of partial response modulation formats compared to NRZ (duobinary = $1 + z^{-1}$*); cumulative PSD is normalized to total signal power in each case.*

nation of the channel and the FIR pre-emphasis filter [11, 18]. The complex data spectrum originating in the transmitter is reshaped such that the resulting waveform available at the receiver after traveling through the channel is a duobinary signal. The transmission system, shown in Figure 5.5, has several main components: a binary data source, a duobinary precoder, a signal spectrum reshaping filter, the channel, a duobinary-to-binary data converter, and a NRZ receiver.
The typical channel will have a frequency roll-off that is much steeper than that of the desired duobinary signal. As a result, the reshaping filter is required to emphasize the higher frequency components as well as to flatten the group delay response across the signal spectrum. As the duobinary data spectrum has a null at half the bit rate, the amount of high frequency emphasis is greatly reduced when compared to uncoded NRZ signaling. Additionally, nulls that occur in the transfer function of the channel are predominantly located towards the higher end of the frequency spectrum; therefore, the compact spectrum of duobinary provides a distinct advantage. The FIR filter used for pre-emphasis is indeed an indispensable element of the design as it allows shaping the signal response towards the desired duobinary shape while respecting the variations in specific channel characteristics. If we define the complex transmission frequency response of the

backplane as $H_{CH}(\omega)$, then the required filter response $H_{FIR}(\omega)$ becomes:

$$H_{FIR}(\omega) = \frac{H_D(\omega)}{H_{CH}(\omega)} \tag{5.1}$$

where $H_D(\omega) = 1 + e^{-j\omega T}$, which is the frequency response of a duobinary filter. In general, an ideal $H_{FIR}(\omega)$ filter has many coefficients. A practical filter will be truncated to only a few filter taps, which are required to suppress the pre-cursor and 2 or 3 post-cursor symbols.

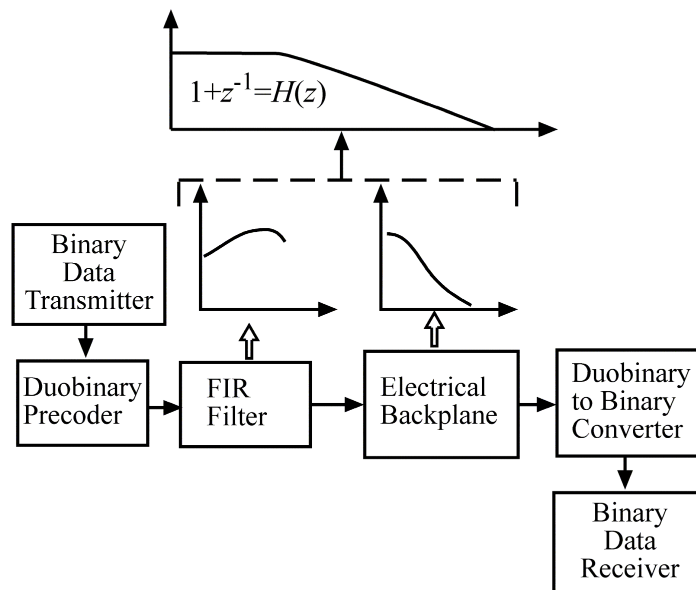


Figure 5.5: Duobinary transmission system architecture.

The low-pass characteristic of duobinary partial response signaling goes beyond the compression of spectrum within the low-loss region of the channel. The limited bandwidth requirement is also beneficial to the design of the front end components of the transceiver. These components can be considered as part of the channel response. This effectively relaxes the integrated circuit bandwidth requirements, opening possibilities to use a broader range of silicon processes and relaxed impedance matching requirements. Similar considerations apply to channel design, such as the mitigation of differential skew. Differential skew converts higher frequency components to common mode, resulting in a low-pass characteristic for the differential mode. This low-pass characteristic can become a part of the duobinary response, relaxing the design criteria even for the highest symbol rates.

A final point to consider, when comparing duobinary to PAM4, is that the redundant information that exists in duobinary signaling is not actually used in the detection process outlined so far. There is, in fact, additional information that can be extracted from limited permissible data transitions. Error detection is briefly discussed by Pasupathy in [9]. It is very possible that some limited error detection can be implemented that would further improve the BER performance of duobinary signaling.

5.3 Custom ASIC design for 50+ Gb/s duobinary link

To support next generation serial 56 Gb/s transfer rates across a backplane there are no off-the-shelf components available. For this speed custom transmitter and receiver chips were designed. The transmitter consists of a feed-forward equalizer (FFE) that shapes the transmission channel (backplane + extra loss in connecting cables etc.) to a duobinary shape. The receiver translates the duobinary input data to 4 quarter rate NRZ streams, using a novel architecture, which is demonstrated to work up to 56 Gb/s.

5.3.1 Feed-forward equalizer

5.3.1.1 Introduction to feed-forward equalization

Feed-forward equalization (FFE) is one of the most common equalization techniques used in serial data paths. Generally, the FFE equalizes the signal by summing the levels from multiple gain-controlled taps representing the weight of the preceding and following voltage level samples. The summation is continuous over the entire waveform.
Compared to other equalization techniques such as decision-feedback equalization (DFE), FFE equalization techniques only correct voltage levels of the received waveform with information about the analog waveform itself. Therefore, the chip design is less complicated and requires fewer gates, thus, in most cases the chip designed using FFE is less expensive and more power efficient.

5.3.1.2 Implementation 12.4 ps spaced 5-tap FFE

Figure 5.6 shows the topology of the 5-tap FFE. The gain cell is a variable gain amplifier. It is implemented as a Gilbert cell, and is the most critical sub-block in the FFE design. These cells realize the equalization coefficients or tap weights. Therefore, the gain stage can be considered as an analog multiplier with a high-speed data input and a low speed control sig-

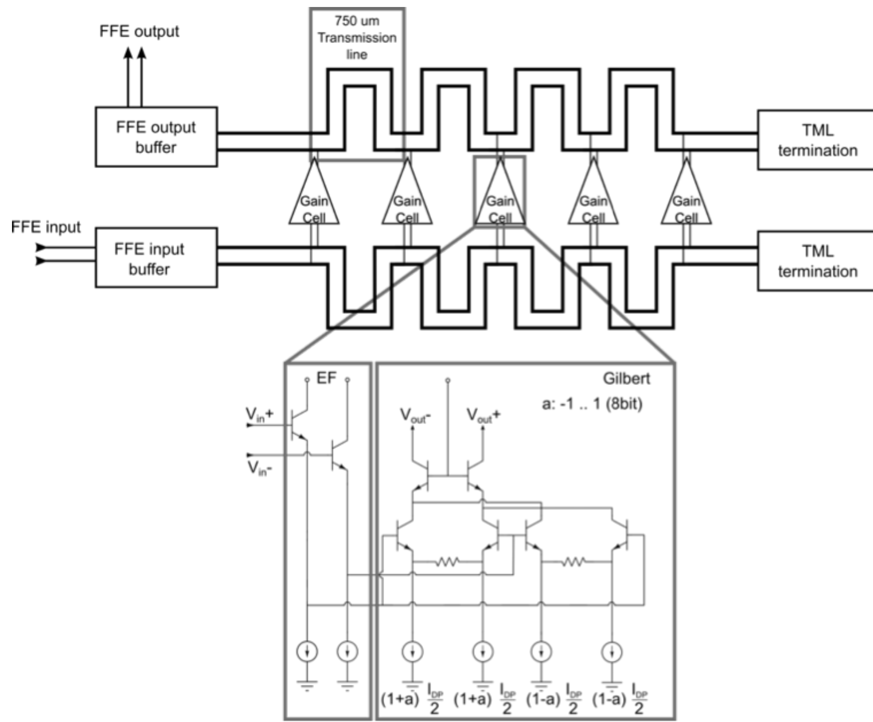


Figure 5.6: 5-tap FFE block diagram.

nal. In addition, by keeping the summed current of both differential pairs constant, the current flowing through the transmission line termination resistor is constant. In this way the bias voltage of the FFE output buffer is kept constant.

As shown in Figure 5.6, the delay of each tap is implemented by high impedance sensing of a transmission line at the input of the gain cells. At the output a high impedance addition on the transmission line is performed. The overall delay of each tap is defined as the sum of the delay at the input and at the output.

On-chip transmission lines (TML) have been used in various FFEs as low-loss delay elements, because of their very high bandwidth and low power dissipation [19]. In this FFE design, each meandered transmission line section in between the gain cells is 750 um long and is designed to have a $50\,\Omega$ characteristic impedance. Meanwhile, the input and output TML are terminated by on-chip resistors.

5.3.1.3 Parameter optimization

Finding the optimal parameters for a duobinary channel in a 5 dimensional space is challenging. However, the methodology proposed in this section provides a good starting point. It is based on frequency domain measurements which can be done fast and accurately. The response of each tap is measured at maximum gain (5 measurements). These measurements are converted into the time domain by calculating the impulse response. Figure 5.7 shows that the 5 taps are separated in time by a delay of 12.4 ps, corresponding to the delays introduced by the transmission lines. Also one can see that the later taps have a lower output power, which is caused by the loss across the transmission line.

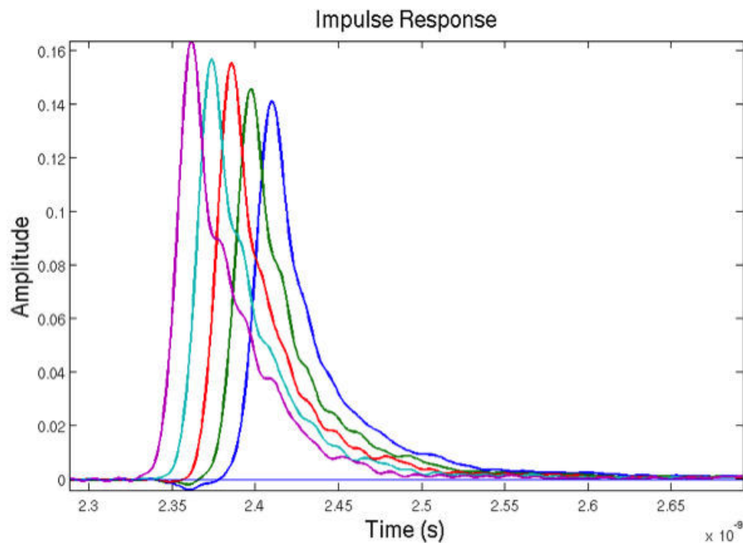


Figure 5.7: Overlapping impulse responses of the FFE.

From the Gilbert cell implementation one can assume the gain of the taps to be linear. As a result, the FFE output can be calculated as a linear combination of the impulse responses of the taps. The complete system can be modeled by the convolution of the channel impulse response and the taps or by multiplying them in the frequency domain and recalculating the impulse responses.
Using the measured impulse responses, the coefficients are fitted using a least square error (LSE) approach to the idealized duobinary response. The idealized response consists of two bit-spaced narrow Sinc pulses. The LSE fit matches the 5 normalized FFE parameters as well as the optimal timing.

In this way it selects the optimal number of pre- and post-cursors in the FFE. The result of such a fit is shown in Figure 5.8.

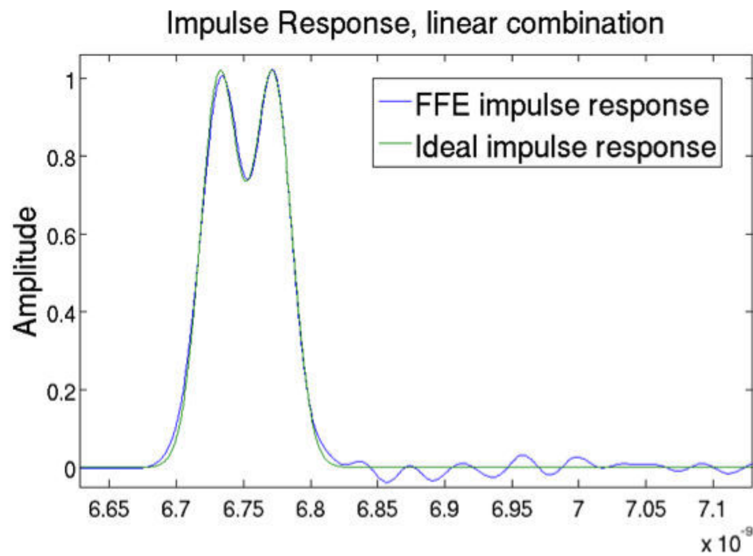


Figure 5.8: Fitting the FFE output to the ideal duobinary response.

It is clear that the FFE is capable of matching the main cursors of the duobinary channel. There is some remaining error which will result in undesired, but acceptable, inter symbol interference, which is evident from the eye diagram of Figure 5.9.

5.3.2 Duobinary receiver

5.3.2.1 Introduction to duobinary receivers

The first duobinary to binary converter proposed by Lender [8] comprised a full wave rectifier. Although this is a viable solution, it is not trivial to scale it up to the multi-gigabit per second range. However, in 2005 Sinsky et al. [11] demonstrated an innovative pseudo digital approach, shown in Figure 5.10, which could potentially be very fast.
To overcome the speed limitation of this duobinary receiver, an on chip demultiplexing step is added before the XOR operation. This architecture is shown in Figure 5.11. To sample the data before decoding, the duobinary signal introduces some extra challenges because the data is highly unbalanced (the ratio of 0's to 1's is 75% compared to the 50% expected when receiving NRZ). Implementing this technique in a fast SiGe BiCMOS pro-

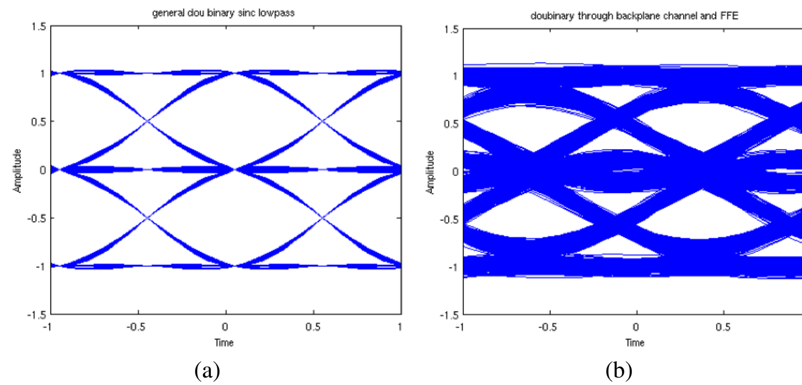


Figure 5.9: Simulated eye diagram of the ideal (left) and FFE-synthesized (right) duobinary signal.

cess allowed us to reach the record breaking speeds of 56 Gb/s across a backplane.

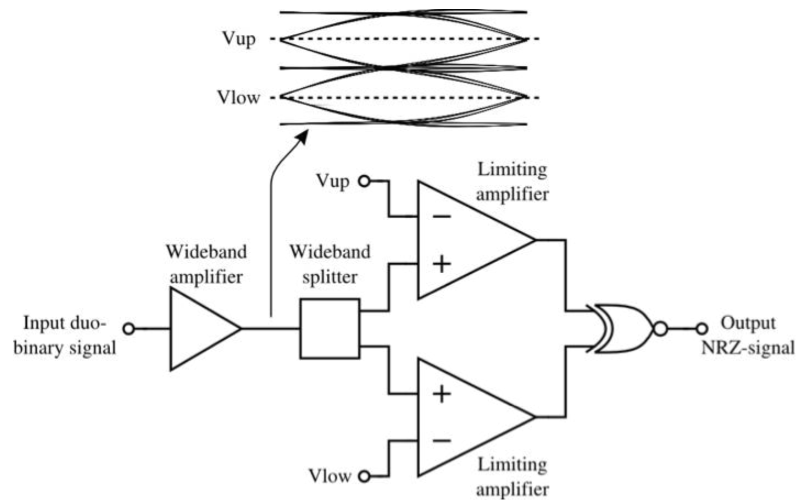


Figure 5.10: First proposed pseudo digital duobinary to binary converter.

After the duobinary decoding, the data is again de-multiplexed and re-timed to reduce the output bit rate. The final data output is a quarter rate stream.

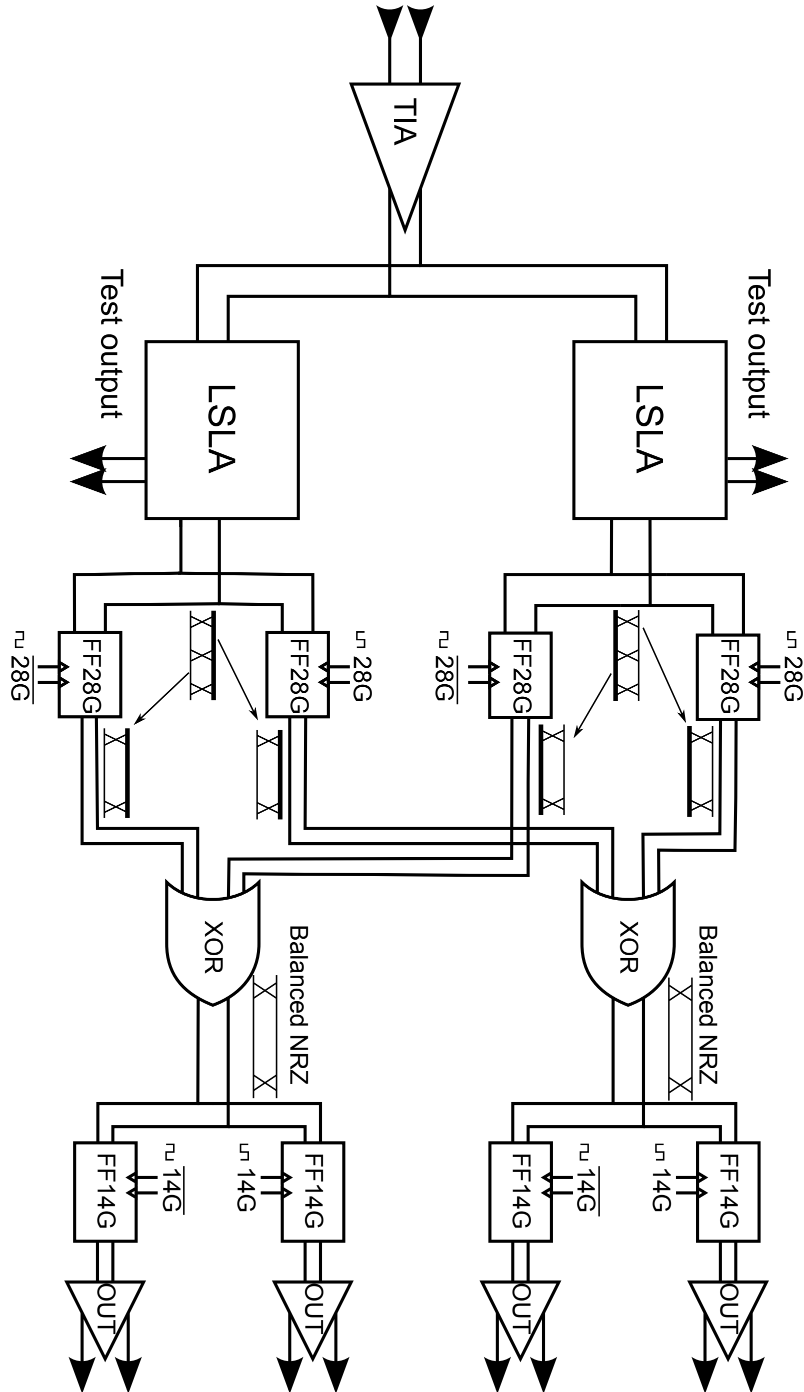


Figure 5.11: Implemented high-speed receiver architecture.

5.4 Eye-pattern and BER measurements

5.4.1 Measurement setup

To evaluate and verify the system, test boards were designed. The TX and RX chips are flip-chip mounted onto these test boards. The data generator is connected to the TX board using coax cables. All coax cables used in the measurement setup are 20 cm long. Each of the coax cables adds 1.05 dB of loss at 28 GHz to the total link loss.

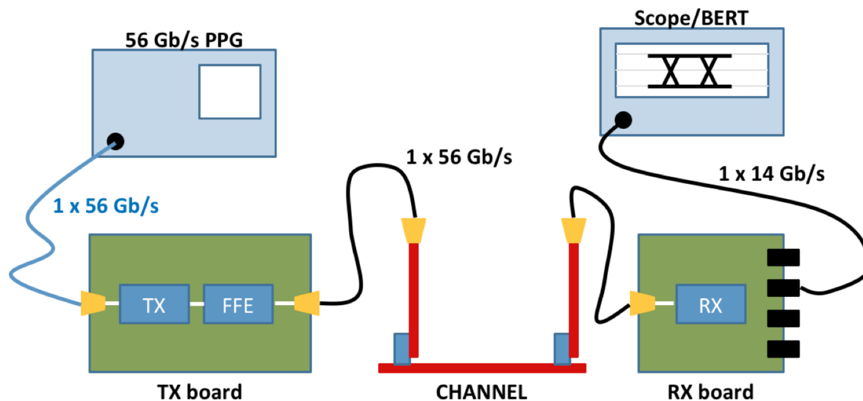


Figure 5.12: Measurement setup.

The complete measurement setup is shown in Figure 5.12. The signal goes through an FFE which pre-shapes the frequency content of the signal at the output of the TX. Figure 5.13 shows the 5.6 dB of losses added by the TX board at 28 GHz.
The output of the TX board is connected to the input of the channel using a pair of coax cables. The output of the channel is connected to the input of the RX board using a second pair of coax cables. The RX board adds an extra 3.8 dB of loss at 28 GHz, as shown in Figure 5.14. Finally, the output of the RX boards is connected to a scope/BERT. The losses added by the different components in the measurement setup are summarized in table 5.1. At 28 GHz a total loss of 11.5 dB is added by the coax cables, the TX board and the RX board. Back-to-back measurements show a vertical and a horizontal eye-opening at the input of the RX board of 57 mV and 12 ps (0.67 UI) respectively. This results in error-free (BER $< 10^{-12}$ without FEC) transmission at 56 Gb/s, the eye-diagram is shown in Figure 5.15.

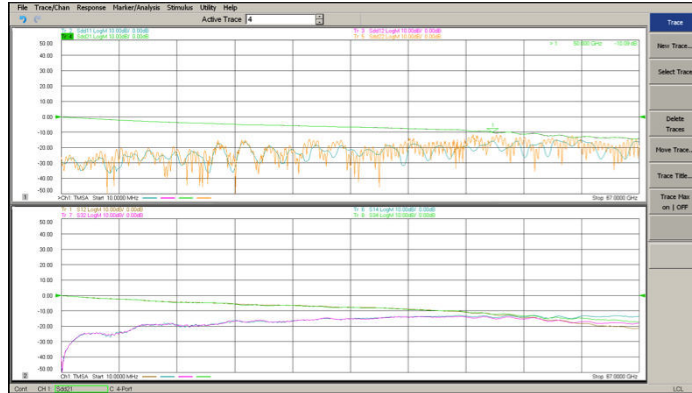


Figure 5.13: TX board losses.

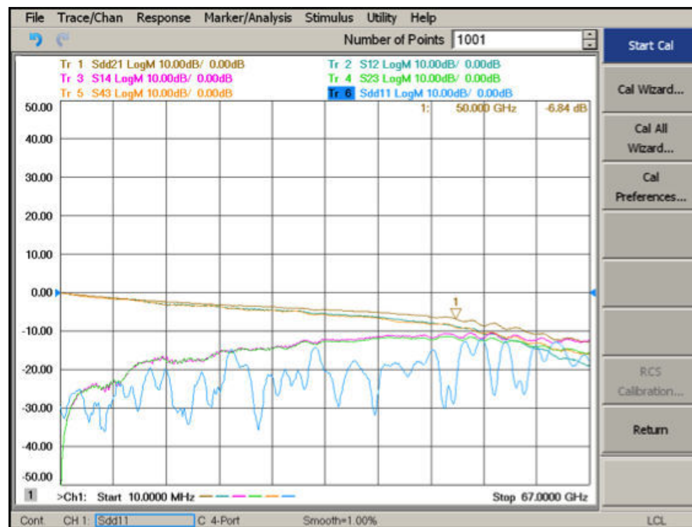


Figure 5.14: RX board losses.

5.4.2 Measurements on ExaMAX® demonstrator

To validate the chip design and to demonstrate 56 Gb/s duobinary transmission over a backplane, measurements have been carried out on a demonstrator using the state-of-the-art ExaMAX® connector system (see Figure 5.15).

The demonstrator consists of 2 daughter cards plugged into a backplane using 2 ExaMAX® connectors. The backplane has 24 layers and is 160

COMPONENT	LOSSES (at 28 GHz)
TX board	5.60 dB
Coax TX board to channel	1.05 dB
Channel losses	IL (dB)
Coax channel to RX board	1.05 dB
RX board	3.8 dB
Total	IL (dB) + 11.5 dB

Table 5.1: Total amount of losses added by measurement environment.

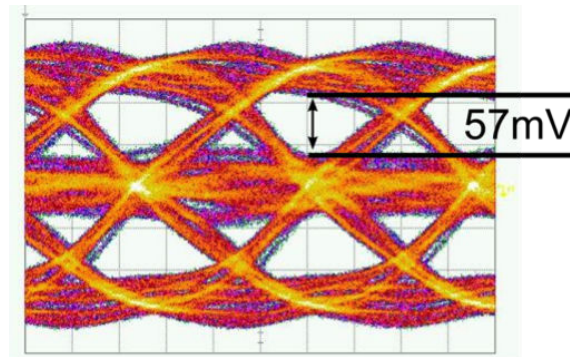


Figure 5.15: 56 Gb/s output eye-diagram of the transmitter.

mil (4.1 mm) thick. The daughter cards have 18 layers and are 94 mil (2.4 mm) thick. The trace lengths on the backplane vary between 1.7 in and 26.75 in. The trace length on the daughter cards is 6 in. This results in a minimum total interconnection length of 13.7 inch (35 cm) + 2 connectors. The material used for building the backplane and daughter cards is Megtron 6. The backplane traces have a loss of 1.3 dB per inch at 28 GHz. The 28 dB insertion loss at 28 GHz of the 13.7 inch channel is shown as the red line in Figure 5.17. As explained above, the measurement setup adds an additional 11.5 dB of loss at 28 GHz. The total loss of the channel including the measurement setup is shown as the blue line in Figure 5.17.
Measurements started at 40 Gb/s on the shortest link (13.7 inch, 35 cm). The loss at 20 GHz is 28.8 dB including measurement setup, resulting in a vertical and a horizontal eye-opening of 18.2 mV and 15 ps (0.6 UI) respectively, while the maximum output eye-pattern at the transmitter has an eye-opening of 93.4 mV and 19.1 ps (0.76 UI) at 40 Gb/s. Both eye-patterns are shown in Figure 5.18. This results in error-free (BER $< 10^{-12}$) transmission when connected to the duobinary decoder.

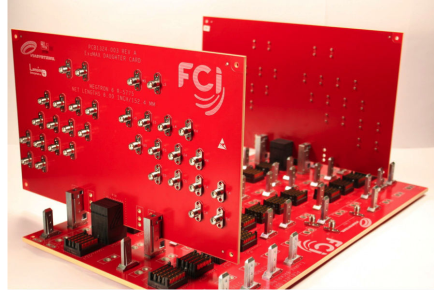


Figure 5.16: ExaMAX® backplane demonstrator.

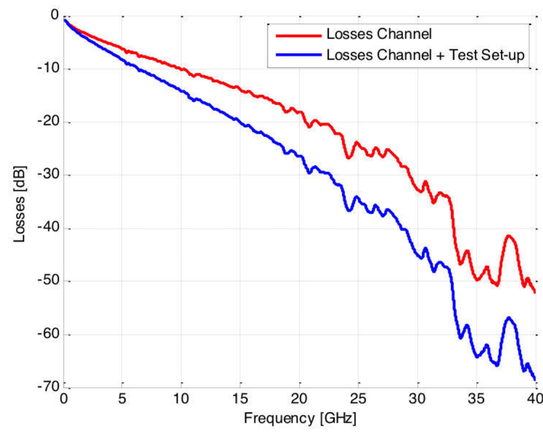
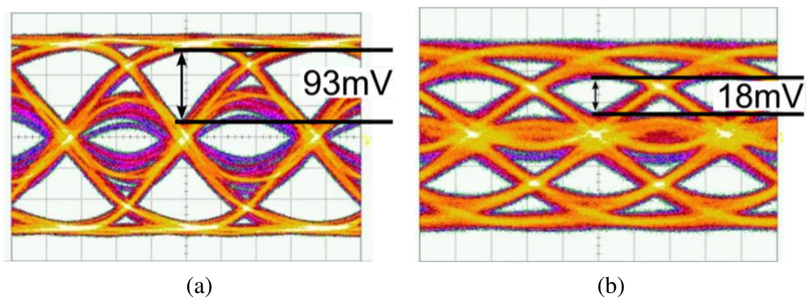


Figure 5.17: Insertion losses 13.7 inch backplane channel only (red) and total losses of backplane channel + test setup (blue).



(a) (b)

Figure 5.18: 40 Gb/s output eye-pattern at the transmitter (left) and after a 13.7 inch backplane channel (right).

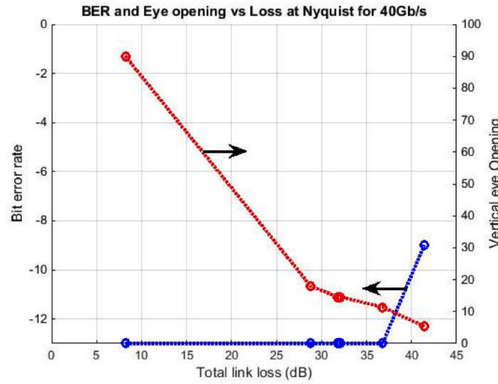
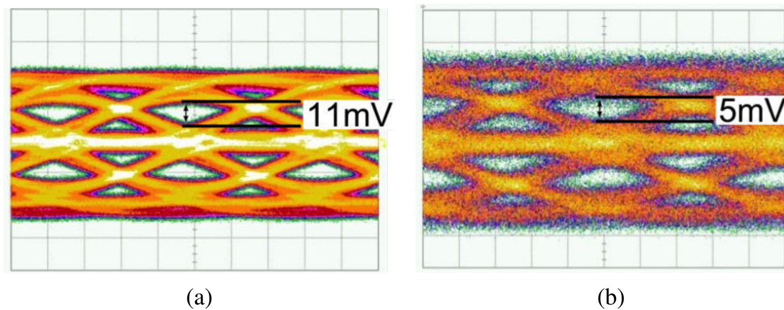


Figure 5.19: BER (blue) and vertical eye-opening (red) as a function of the loss at the Nyquist frequency for a 40 Gb/s signal measured across the ExaMAX® backplane.

In Figure 5.19 it is shown that a total loss (backplane including test setup) at 20 GHz of up to 36.8 dB can be equalized to result in error-free transmission (BER $< 10^{-12}$), and up to 41.4 dB with a BER below $5 \cdot 10^{-9}$, which is considered OK for a link with FEC in the current 25 Gb/s IEEE 802.3bj standard. The 36.8 dB loss corresponds to a total channel length of 22 in, while the 41.4 dB loss corresponds to a total channel length of 27 in. At 36.8 dB and 41.4 dB the vertical eye-opening are 11 mV and 5 mV respectively, as shown in Figure 5.20.

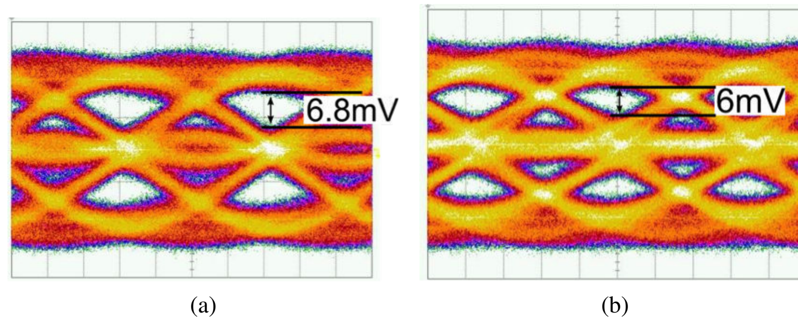


(a) (b)

Figure 5.20: Eye-diagrams of the 40 Gb/s duobinary signal across a 22 inch (left) and a 27 inch (right) backplane channel.

Moving towards higher speeds leads to more frequency-dependent loss. At 50 Gb/s the signal after a 13.7 inch backplane channel was still received error-free (BER $< 10^{-12}$), with an eye-opening of 6.8 mV as shown on

the left in Figure 5.21.
By increasing the speed to 56 Gb/s, the loss at the Nyquist frequency increases further, and the vertical eye-opening at the input of the receiver decreases to about 6 mV, shown on the right in Figure 5.21. The BER obtained at 56 Gb/s is $< 5 \cdot 10^{-9}$, which is more than sufficient assuming FEC is applied.



(a) (b)

Figure 5.21: Eye-diagrams of the 50 Gb/s (left) and 56 Gb/s (right) signal after a 13.7 inch backplane channel.

5.5 Measurements on active daughter cards

The performance of the transceiver chipset is limited by the total amount of loss that can be compensated by the FFE. In Section 5.4.1, it was shown that 11.5 dB of the losses at 28 GHz are added by the measurement setup. These losses are caused by the testboards and the coax cables needed to connect the testboards to the backplane demonstrator. They can be drastically reduced by designing active daughter cards that are plugged directly on to the backplane demonstrator. The design of these daughter cards was explained in detail in Section 3.6.

Because the chips are now directly flip-chip connected to the channel, it is not possible to measure the exact insertion loss (IL) of a link. However, there are test structures on the daughter cards that allow to either input cross talk aggressors or verify the channel performance including ExaMAX® footprint. The measurement of 2 different channel lengths is shown in Figure 5.22. The dark grey channel is about 57 cm (22.5 inch) and has a loss at 28 GHz of 44.4 dB, the grey channel is 40 cm with an IL of 33.3 dB. During the routing of the daughter card, it was not possible

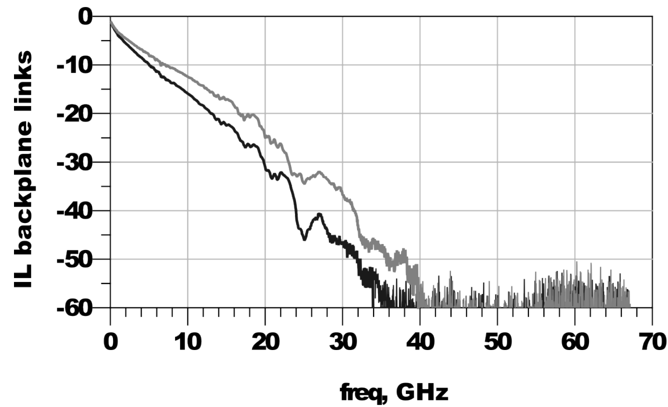
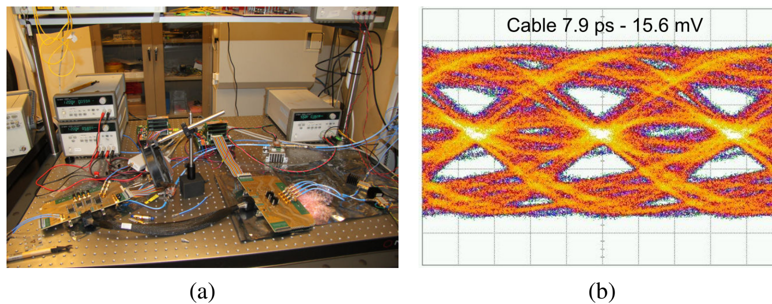


Figure 5.22: Link loss measured through 1.85 mm channels. The grey graph has 33.3 dB of loss at 28 GHz and is about 40 cm, the dark grey graph has about 44.4 dB of loss at 28 GHz and is about 57 cm.

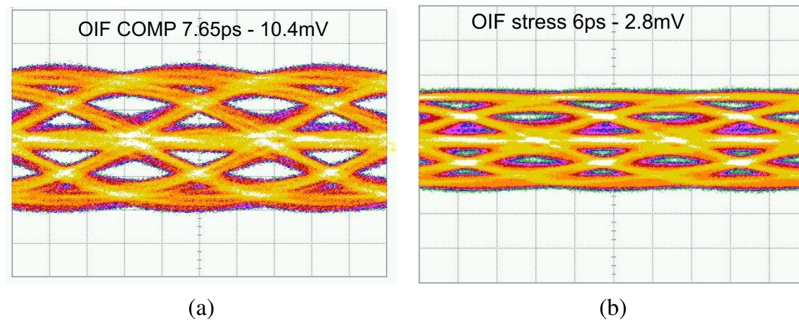
to have the exact same length for the transmitter output, the receiver input and the 1.85 mm channels. This results in the chip-to-chip link on the short channel having a length of 33 cm (13 inch) and the long channel having a length of about 50 cm (20 inch). Due to the shorter link length (about 7.5 cm shorter), the resulting loss will also reduce with 4 dB (1.3 dB/inch as measured in Section 3.6).



(a) (b)

Figure 5.23: First time domain measurments on active daughter cards with short cable backplane.

Measurements verify error-free transmission up to about 20 inch consisting of 2.7 inch at receiver side, 14.25 inch on the backplane and 2.3 inch at transmitter side, with a total measured link loss of 40.4 dB at 28 GHz and a BER of about 10^{-12} at 50 Gb/s with a link length of 66 cm (loss of about 49 dB at 28 GHz, 20.5 inch on the backplane).



(a) (b)

Figure 5.24: 56 Gb/s eye diagrams across the active backplane setup wih a 40 cm link on the left and a 57 cm link on the right.

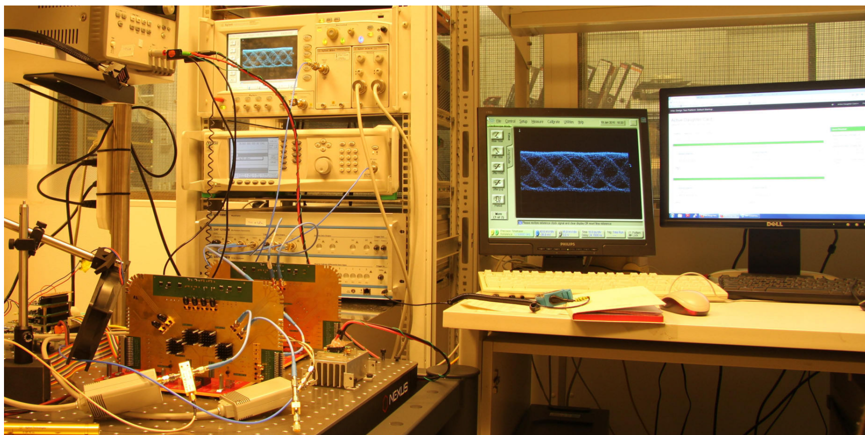


Figure 5.25: Time domain measurement setup with a 20 inch backplane link (50 cm).

The active daughter cards allow us to verify the maximum link length and associated loss, while adding the possibility to measure crosstalk. No quantitative crosstalk measurements are yet performed, however, some qualitative measurements were done by launching a 28 Gb/s duobinary modulated stream as a far end crosstalk aggressor. This did not influence the error-free reception of the 56 Gb/s data. The measurement with 28 Gb/s aggressor was demonstrated live at DesignCon 2015, as shown in Figure 5.27.

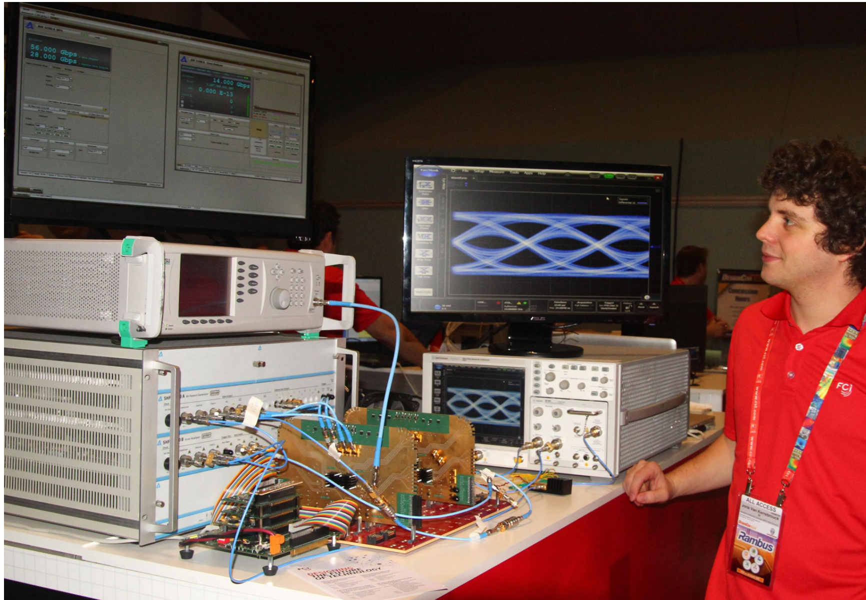


Figure 5.26: Error-free transmission with far end crosstalk agressor demonstrated live at DesignCon 2015.

5.6 Conclusions

In this chapter, it is shown that 56 Gb/s transmission across commercial backplane connectors and with available chip technologies is possible using duobinary signaling and an FFE consisting of only 5-taps. Initial measurements have shown that it is possible to transmit 56 Gb/s duobinary signals successfully over a channel with up to 40 dB of loss at 28 GHz. Mounting the chips directly on Megtron 6 daughtercards results in error-free transmission across 20 inches (40 dB insertion loss at 28 GHz). The paper presented in this chapter won a best paper award at DesignCon 2015.

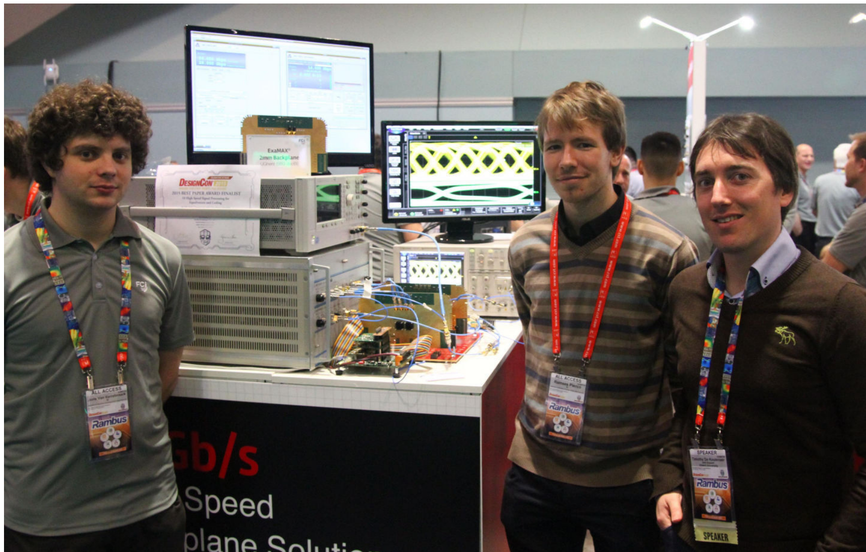


Figure 5.27: The team at DesignCon 2015, Joris on the left, Ramses in the middle and myself on the right.

References

[1] J.H. Sinsky, A. Gnauck, B. Kozicki, S. Sercu, A. Konczykowska, A. Adamiecki, and M. Kossey. 42.8 Gb/s pam-4 data transmission over low-loss electrical backplane. *Electronics letters*, 48(19):1206–1208, 2012.

[2] D.G. Kam, TJ. Beukema, Y.H. Kwark, L. Shan, X. Gu, P.K. Pepeljugoski, and M.B. Ritter. Multi-level signaling in high-density, high-speed electrical links. *Proc. IEC DesignCon (Santa Clara, CA, 2008)*, 2008.

[3] J. Maes, M. Guenach, K. Hooghe, and M. Timmers. Pushing the limits of copper: Paving the road to ftth. In *2012 IEEE International Conference on Communications (ICC)*, pages 3149–3153. IEEE, 2012.

[4] T. Tanaka, T. Takahara, H. Isono, H. Sakamoto, and Y. Miyaki. Dependency of transmission parameters for 400GbE dmt 10km transceiver. *IEEE 802.3bs May 2014*, 2014.

[5] T. Beukema. Design considerations for high-data-rate chip interconnect systems. *IEEE Communications Magazine*, 48(10):174–183, 2010.

[6] R. Kollipara and B. Kirk. Practical design considerations for 10 to 25 Gb/s copper backplane serial links. 2006.

[7] A. Singh, D. Carnelli, A. Falay, K. Hofstra, F. Licciardello, Kia S., H. Santos, A. Shokrollahi, R. Ulrich, C. Walter, et al. 26.3 a pin-and power-efficient low-latency 8-to-12 Gb/s/wire 8b8w-coded serdes link for high-loss channels in 40nm technology. In *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 442–443. IEEE, 2014.

[8] A. Lender. The duobinary technique for high-speed data transmission. *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics*, 82(2):214–218, 1963.

[9] S. Pasupathy. Correlative coding: a bandwidth-efficient signaling scheme. *IEEE Communications Society Magazine*, 15(4):4–11, 1977.

[10] P. Kabal and S. Pasupathy. Partial-response signaling. *IEEE Transactions on Communications*, 23(9):921–934, 1975.

[11] D.G. Kam, TJ. Beukema, Y.H. Kwark, L. Shan, X. Gu, P.K. Pepeljugoski, and M.B. Ritter. Multi-level signaling in high-density, high-speed electrical links. *Proc. IEC DesignCon (Santa Clara, CA, 2008)*, 2008.

[12] J.H. Sinsky, A. Konczykowska, A. Adamiecki, F. Jorge, and M. Duelk. 39.4 Gb/s duobinary transmission over 24.4 m of coaxial cable using a custom indium phosphide duobinary-to-binary converter integrated circuit. *IEEE Transactions on Microwave Theory and Techniques*, 56(12):3162–3169, 2008.

[13] J. Lee, N. Kaneda, T. Pfau, A. Konczykowska, F. Jorge, J.Y. Dupuy, and Y.K. Chen. Serial 103.125 Gb/s transmission over 1 km ssmf for low-cost, short-reach optical interconnects. In *Optical Fiber Communication Conference*, pages Th5A–5. Optical Society of America, 2014.

[14] I. Lyubomirsky. Quadrature duobinary for high-spectral efficiency 100g transmission. *Journal of Lightwave Technology*, 28(1):91–96, 2010.

[15] K. Yonenaga and S. Kuwano. Dispersion-tolerant optical transmission system using duobinary transmitter and binary receiver. *Journal of Lightwave Technology*, 15(8):1530–1537, 1997.

[16] D. van Veen, V. Houtsma, A. Gnauck, and P. Iannone. 40 Gb/s tdm-pon over 42 km with 64-way power split using a binary direct detection receiver. In *2014 European Conference on Optical Communication (ECOC)*, pages 1–3. IEEE, 2014.

[17] D. van Veen, V.E. Houtsma, P. Winzer, and P. Vetter. 26 Gb/s pon transmission over 40-km using duobinary detection with a low cost 7-GHz APD-based receiver. In *2012 European Conference on Optical Communication (ECOC)*, pages 1–3. IEEE, 2012.

[18] W.T. Beyene and A. Amirkhany. Controlled intersymbol interference design techniques of conventional interconnect systems for data rates beyond 20 Gb/s. *IEEE Transactions on Advanced Packaging*, 31(4):731–740, 2008.

[19] A. Momtaz and M.M. Green. An 80 mW 40 Gb/s 7-tap t/2-spaced feed-forward equalizer in 65nm cmos. *IEEE Journal of Solid-State Circuits*, 45(3):629–639, 2010.

[20] T. De Keulenaer, Y. Ban, Z. Li, and J. Bauwelinck. Design of a 80 Gb/s SiGe bicmos fully differential input buffer for serial electrical communication. In *2012 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 237–239. IEEE, 2012.

6
Conclusions and future research

IN this dissertation a high-speed duobinary receiver architecture is described. The implementation of this architecture worked at data rates above 80 Gb/s. Testing at these rates becomes increasingly difficult. Hence, to facilitate system tests, board design methodologies were developed and different materials and connectors were characterized. The combination of the designed chips mounted on the test boards gave the opportunity to achieve an impressive high-speed system and the accompanying test results. This led to three patent applications and multiple papers in international journals and conference proceedings on chip implementation, high-speed board design and high-speed link tests.

This final chapter summarizes the results and the possibilities of the realized chipset along with the paths for valorization and some ideas on further research in the field of high-speed serial communication.

6.1 Summary of the results

In this work the three main modulation schemes for high-speed serial links were compared and duobinary modulation was selected as the most promising for backplane communication. The selection of duobinary modulation is based on the trade off between spectral efficiency, implementation complexity and power consumption, as was illustrated in Chapter 1.

In Chapter 2, the main focus was on the chip design and a high-speed duobinary receiver was designed with a matching bandwidth of 50 GHz capable of receiving duobinary signals up to 100 Gb/s. The complete receiver includes a duobinary front end and a DEMUX/XOR. The simulated power consumption was roughly 1.4 W resulting in 14 pJ/bit for 100 Gb/s (including all test outputs). The research described in this chapter led to two conference publications [4, 5] and two filed patents (EP14161772.0 and EP14305284.3).

In Chapter 3, the importance of board design and how to mount a high-speed chip on a board was highlighted. The design of connectorized traces was shown to have a smooth frequency response up to 67 GHz. Different connectors were compared and to achieve the best performance the 1.85 mm Rosenberger connector was selected because of its low loss and the ability to place it in the middle of a board. A 50 GHz rat-race was designed to have excellent phase and amplitude matching. Using these techniques a test board for the high-speed receiver chip was designed, demonstrating the receivers chip's performance at 84 Gb/s. Next to the 4 layer test board, an advanced 12 layer active Megtron 6 daughter card was designed. On this board, four chips were mounted, two receivers and two transmitters, with all four routed to the FCI ExaMAX® connector. This board was used together with a Megtron 6 ExaMAX® backplane demonstrator. This chapter was based on a publication presented at the IEEE Workshop on Signal and Power Integrity 2014 [6].

In Chapter 4, the receiver chip was demonstrated to work at 84 Gb/s. A 84 Gb/s transmission link is achieved across 20 cm of coax cable connected in between the receiver and transmitter test boards with a BER $< 10^{-11}$. The total channel loss at 42 GHz is approximately 14 dB with the channel equalized to the duobinary shape by the 5 tap feedforward equalizer available at the transmitter. On the date of publishing, this was the fastest duobinary modulated serial communication link presented in a journal paper published in Electronic Letters [2].

Finally, a complete backplane test bed was built to show the chips performance at 56 Gb/s. The test bed contains two active daughter cards, each with 4 chips mounted on them, and a Megtron 6 ExaMAX® backplane. Error free (BER $< 10^{-13}$) transmission was shown across backplane links up to 50 cm with loss in excess of 40 dB at 28 GHz. Transmission across a channel with 33 dB of loss at 28 GHz was demonstrated live at DesignCon 2015 in Santa Clara (USA), in the booth of FCI with a 28 Gb/s cross talk aggressor. A paper on this work was also presented at DesignCon 2015 and

won a best paper award [3].

Besides the electrical measurement setups built with the duobinary receiver, there are also some ongoing optical experiments. The first of these was presented at the Optical Fiber Communication Conference (OFC) 2015 [1], and we are invited by the Journal of Lightwave Technology to prepare a more extensive journal paper on this topic.

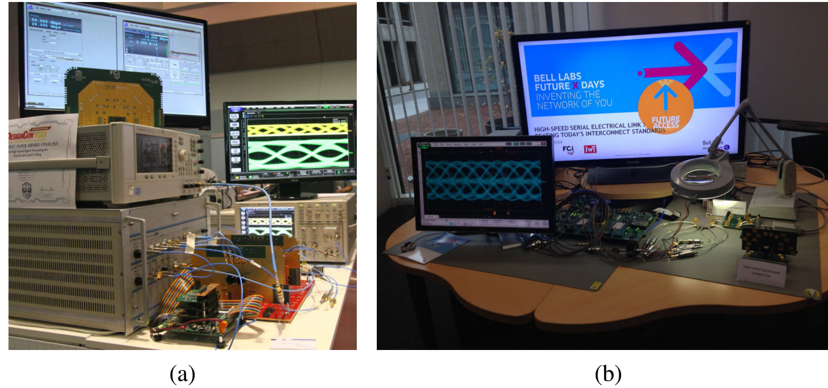
6.2 Valorisation opportunities

The work presented in this dissertation gained a lot of attention from the industry during the Bell Labs Future X days demos as well as at DesignCon 2015. Photos of the demos are given in Figure 6.1. The developed chipset and board design techniques can be used for many different applications, such as: driving long high-speed optical channels, transmission across complex high-loss backplane channels or short-range chip-to-module applications.

The lab was contacted on numerous occasions to demonstrate our high data rate duobinary signaling solution across multiple connector solutions. System vendors also contacted us for a demo across long PCB traces with 40 dB of loss, which all lies within the capabilities of this chipset.

At the Bell Labs Future X days demo, where 84 Gb/s duobinary signaling was demonstrated, other interesting parties also looked into the possibilities of this chipset. E.g. Barco saw an application for reducing the amount of cables for transporting uncompressed 8K video.

Patent applications were filed during the design process to protect this unique IP and support the start of a spin-off company commercializing this chipset. For this, the addition of a clock and data recovery circuit is one of the most crucial steps.



(a) (b)

Figure 6.1: Demo of 56 Gb/s backplane comunnication at DesignCon 2015 (Santa Clara) (a) and 84 Gb/s duobinary signaling demo at the Bell Labs Future X days (Antwerp) (b).

6.3 Future research

Besides the successful implementation of the complete transceiver, showing the feasibility of single lane 100 Gb/s interconnects, other system components are required to build a complete standalone system. This opens several paths to follow-up projects and new research challenges.
Among others, a robust system needs different clock and data recovery (CDR) circuits, both at the input of the transmitter and at the receiver side, making sure the phase of the DEMUX clock is optimally aligned.
Next to a CDR circuit, automated tap weight calibration would facilitate the complete setup of the shaping filter as well as the eye threshold values. To achieve this, bidirectional communication is needed to indicate the signal quality at the receiver side as well as automatic high-speed test pattern generation to test the link quality.

There are not only improvements possible on an implementation level. The mounting of this chip during the test phases was done using direct flip chip on board, which will not be cost effective in a final product. This is the subject of a whole other field of research looking for 100 Gb/s packaging.

Clearly, this work contributes greatly to the advancement of the state-of-the-art in the field of high-speed transceivers, although to elevate it to the next level and make it a commercially viable, user-friendly system requires fine tuning and a lot of dedication.

References

[1] X. Yin, J. Verbist, T. De Keulenaer, B. Moeneclaey, J. Verbrugghe, XZ. Qiu, and J. Bauwelinck. 25 Gb/s 3-level burst-mode receiver for high serial rate TDM-PONs. In *Optical Fiber Communication Conference*, paper Th4H.2. Optical Society of America, 2015.

[2] T. De Keulenaer, G. Torfs, Y. Ban, R. Pierco, R. Vaernewyck, A. Vyncke, Z Li, J.H. Sinsky, B. Kozicki, X. Yin, and J. Bauwelinck. 84 Gb/s SiGe bicmos duobinary serial data link including serialiser/deserialiser (serdes) and 5-tap ffe. *Electronics Letters*, 51(4):343–345, 2015.

[3] T. De Keulenaer, J. De Geest, G. Torfs, J. Bauwelinck, Y. Ban, J.H. Sinsky, and B. Kozicki. 56+ Gb/s serial transmission using duobinary signaling. *Proc. IEC DesignCon (Santa Clara, CA, 2015)*, 2015.

[4] T. De Keulenaer, G. Torfs, R. Pierco, and J. Bauwelinck. A digitally controlled threshold adjustment circuit in a 0.13μm SiGe bicmos technology for receiving multilevel signals up to 80 Gb/s. In *2014 10th Conference on Ph. D. Research in Microelectronics and Electronics (PRIME)*, pages 1–4. IEEE, 2014.

[5] T. De Keulenaer, Y. Ban, Z. Li, and J. Bauwelinck. Design of a 80 Gb/s SiGe bicmos fully differential input buffer for serial electrical communication. In *2012 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 237–239. IEEE, 2012.

[6] T. De Keulenaer, Y. Ban, G. Torfs, S. Sercu, J. De Geest, and J. Bauwelinck. Measurements of millimeter wave test structures for high speed chip testing. In *2014 IEEE 18th Workshop on Signal and Power Integrity (SPI)*, pages 1–4. IEEE, 2014.