Equalization for high-rate communication over optical fiber

Jelle Bailleul, Johannes Van Wonterghem

Supervisors: Prof. dr. ir. Marc Moeneclaey, Prof. dr. ir. Johan Bauwelinc

Counsellors: Dr. ir. Guy Torfs, Dr. ir. Lennert Jacobs, Prof. dr. Xin Yin, Bart Moeneclaey, Ir. Timothy De Keulenaer

Master's dissertation submitted in order to obtain the academic degree of Master of Science in Electrical Engineering

Department of Telecommunications and Information Processing
Chairman: Prof. dr. ir. Herwig Bruneel

Department of Information Technology
Chairman: Prof. dr. ir. Daniël De Zutter

Faculty of Engineering and Architecture
Academic year 2014-2015
Equalization for high-rate communication over optical fiber

Jelle Bailleul, Johannes Van Wonterghem

Supervisors: Prof. dr. ir. Marc Moeneclaey, Prof. dr. ir. Johan Bauwelinc

Counsellors: Dr. ir. Guy Torfs, Dr. ir. Lennert Jacobs, Prof. dr. Xin Yin, Bart Moeneclaey, Ir. Timothy De Keulenaer

Master's dissertation submitted in order to obtain the academic degree of Master of Science in Electrical Engineering

Department of Telecommunications and Information Processing
Chairman: Prof. dr. ir. Herwig Bruneel

Department of Information Technology
Chairman: Prof. dr. ir. Daniël De Zutter

Faculty of Engineering and Architecture
Academic year 2014-2015
Preface

This master thesis is the result of one year of hard work. In this last year we experienced many alternating moments of joy and despair. The challenges we encountered tested our persistence, but in the end the satisfaction of solving them predominates. Along the way, many people have been of importance to the eventual success of this thesis.

To start, we want to thank our supervisors prof. dr. ir. Marc Moeneclaey and prof. dr. ir. Johan Bauwelinck. This thesis would not have been possible without their support. They gave us the freedom to establish our own ideas, but were always accessible if problems arose. We are especially grateful for the discussions with professor Moeneclaey who provided us with important insights.

Further we really appreciated the support that was given by ir. Bart Moeneclaey. He helped us to set up the measurements, showed us the operation of the different measurements instruments and was always available for extra questions. Next we want to thank ir. Joris Van Kerrebrouck who provided us with the S-parameters of the continuous-time equalizer. As the equipment we used is really expensive, we would also like to thank Intec Design, for giving us the opportunity to perform our measurements.

Since we spent a considerable amount of time in the thesis room, we also want to thank our fellow students for the good atmosphere and mutual encouragement to complete this thesis.

Finally we explicitly want to say thanks to our parents, family and friends for the unconditional support they gave us this past year.

Permission for consultation

“The authors give permission to make this master dissertation available for consultation and to copy parts of this master dissertation for personal use. In the case of any other use, the copyright terms have to be respected, in particular with regard to the obligation to state expressly the source when quoting results from this master dissertation.”

Abstract

Dispersion limits the maximal bit rate and fiber length in optical communication systems such as passive optical networks (PONs). Because of the photodetector (PD) non-linearity, the optical fiber chromatic dispersion cannot easily be compensated at the receiver side. We investigate the performance of discrete equalizers based on the minimal mean square error (MMSE) criterion in this situation using computer simulation; linear equalization, linear fractional equalization and (non-linear) decision-feedback equalization are treated. Measurements are also performed on a PON emulation test setup with a continuous time equalizer. We adapt the discrete equalizer algorithms to a practical algorithm for the continuous time equalizer and discuss the results.

Keywords

PON, chromatic dispersion, MMSE, fractional equalizer, continuous time equalizer
Equalization for high-rate communication over optical fiber

Jelle Bailleul, Johannes Van Wonterghem

Supervisor(s): prof. dr. ir. Marc Moeneclaey, prof. dr. ir. Johan Bauwelinck

Abstract—Dispersion limits the maximal bit rate and fiber length in optical communication systems such as passive optical networks (PONs). Because of the photodetector (PD) non-linearity, the optical fiber chromatic dispersion cannot easily be compensated at the receiver side. We investigate the performance of discrete equalizers based on the minimal mean square error (MMSE) criterion in this situation using computer simulation; linear equalization, linear fractional equalization and (non-linear) decision-feedback equalization are treated. Measurements are also performed on a PON emulation test setup with a continuous time equalizer. We adapt the discrete equalizer algorithms to a practical algorithm for the continuous time equalizer and discuss the results.

Keywords— PON, chromatic dispersion, MMSE, fractional equalizer, continuous time equalizer

I. INTRODUCTION

To accommodate the ever-increasing bandwidth demands, network operators have to move the optical-electrical conversion ever closer to their subscribers. Deployment of PONs to offer fiber to the home (FTTH) connections is an important step in this evolution. The typical architecture of a PON consists of an optical line termination (OLT) at the network operator’s central office (CO), a passive optical splitter and a number of optical network units (ONUs) at the premises of the end users. Optical interconnections are provided with single mode optical fibers. The point-to-multipoint topology minimizes the number of optical transceivers and the passivity of the intermediate network nodes means we do not need to provide power alongside the fiber, which makes it low cost, easy to maintain and reliable.

PON standards prescribe an optical loss budget based on the maximum allowable bit error rate (BER), this loss budget limits the amount of subscribers and/or the maximal OLT-ONU distance. In this thesis we study electrical equalization of the received signal to decrease the BER for a given received optical power. Multiple discrete time equalization strategies are simulated and compared and we also treat continuous time equalization. The performance of the latter is also tested in a PON emulation test setup.

II. SYSTEM MODEL

To evaluate and characterize the discrete equalizers we design, we study and simulate a model of the PON communication system and the equalization algorithms. Restriction to only the essential components of the PON leads to the model that is schematically shown in Fig. 1.

At the transmitter, the data bits \(a_k \in \{0, 1\}\) are mapped on a unipolar non-return-to-zero (NRZ) transmit pulse at a rate of \(1/T\). The resulting signal \(e_{in}\) is sent over the optical fiber, modeled by its chromatic dispersion frequency response

\[
H_{fs}(f) = \exp(-j\gamma f^2 T^2)
\]

where

\[
\gamma = \pi D(\lambda) \frac{\lambda^2}{cT^2} L
\]

and \(\lambda\) is the optical wavelength, \(L\) the fiber length and \(D(\lambda)\) the chromatic dispersion parameter depending on \(\lambda\) and the fiber type; we assume \(D(\lambda) = 15 \text{ps/}(\text{nm.km})\).

At the receiver the optical signal with envelope \(e_{out}\) is converted to the current \(i_{PD}\) by the PD and zero-mean stationary white Gaussian noise \(i_n\) with power spectral density (PSD) \(N_0/2\) is added. The resulting current \(i_{PD,n}\) is applied to a transimpedance amplifier (TIA) with second-order low pass frequency response and the voltage at its output is sampled at a rate of \(M_s/T\). These samples \(r(k)\) are applied to an equalizer, decision is carried out on its output \(u\).

The natural pulsation \(\omega_0\) and damping factor \(\zeta\) of the TIA are optimized for the system where \(\gamma = 0\). There the channel and PD do not change the signal and the analytical worst case BER is given by

\[
\text{BER}_{\text{max}} = Q\left(\frac{0.5h_{\text{tot}}(\tau)| - \text{ISI}_{\text{max}}(\tau)}{\sigma_0}\right)
\]

with \(Q(x)\) the tail probability of the standard normal distribution, \(\sigma_0^2 = (N_0/2)(\omega_0/4\zeta)\) the variance of the noise contribution to \(r\) and \(0.5|h_{\text{tot}}(\tau)| - \text{ISI}_{\text{max}}(\tau)\) half the vertical eye opening at time instant \(kT + \tau\) at the TIA output without noise. Optimizing 3 yields \(\omega_0,\zeta = 3.1652/T\) and \(\zeta_{\text{opt}} = 0.6707\) which corresponds to an optical power penalty of 0.39 dB compared to the receive filter matched to the NRZ pulse.

We express the performance of our system in terms of BER curves as function of the signal to noise ratio \(E_b/N_0\). \(N_0/2\) is the power spectral density of the white noise at the TIA input, \(E_b\) is the energy per bit at the output of the fiber, scaled by the PD responsivity, and conveniently lumps the continuous wave (CW) laser power, Mach-Zehnder modulator (MZM) losses, splitter attenuation and fiber attenuation. BER simulations are carried out semi-analytically[3] to limit simulation times.
III. DISCRETE TIME MMSE EQUALIZERS

To compensate the intersymbol interference (ISI) and improve the system performance equalization can be applied to the sample of the received signal \( r \). The MMSE criterion is used to derive the equalizer coefficients because decreasing the mean square error (MSE) generally also decreases the BER and because deriving an expression for the equalizer coefficients is straightforward.

A. Linear Equalizer

A first type of equalizer we consider is the linear equalizer whose output \( u(k) \) is a linear combination of \( N_{FF} \) samples of \( r(k) \) and has \( N_{FF,l} \) non-causal taps, \( N_{FF,r} \) causal taps and an offset \( c_{off} \). The MSE expression becomes

\[
\text{MSE} = E[(u(k) - a(k))^2] = E[\sum_{m} r(k - m)h_{FF}(m) - a(k) - c_{off}]^2
\]

Differentiating this expression to \( h_{FF} \) and \( c_{off} \) and equating to zero yields a linear system of equations for \( h_{FF} \)

\[
Rh_{FF} = r \rightarrow h_{FF} = rR^{-1}
\]

The ISI as a result of the dispersion is not easily compensated because the phase information of \( e_{out} \) is lost after the PD non-linearity. Simulations confirm this as the linear equalizer does not succeed in removing the BER floor that occurs for larger \( \gamma \).

B. Linear Fractionally Spaced Equalizer

We show and verify that a finite length zero forcing (ZF) fractionally spaced equalizer exists for an arbitrary non-linear channel if

\[
J(N_{eq} + N - 1) \leq M_s N_{eq}
\]

where \( N \) is the channel filter length, \( J \) is the number of non-linear contributions types, \( M_s \) is the number of samples per symbol period and \( M_s N_{eq} \) is the number of equalizer coefficients. As the ZF and MMSE equalizer converge to each other at high \( E_b/N_0 \), we also expect the BER floor to disappear if we increase the \( M_s \) and \( N_{eq} \) of our MMSE fractionally spaced equalizer.

C. Decision Feedback Equalizer

If we introduce decision feedback, the equalizer now consists of a feedforward filter \( h_{FF} \) and a feedback filter \( h_{FB} \). This improves the equalizer performance, especially as the feedback filter uses previously estimated symbols and hence does not enhance the noise. The equalizer now is no longer linear, but the MMSE criterion still reduces to a system of equations of the general shape 5. A disadvantage of decision feedback is that error propagation can occur, we verify that this is not catastrophic and that the performance degradation is limited.

D. Non-Linear Decision Feedback Equalizer

As the PD non-linearity introduces quadratic terms in the data symbols, the equalizer performance should increase if we also include these in the feedback filter (i.e. \( h_{FB} \) becomes a Volterra filter).

E. Simulation Results

In Fig. 2 we show the simulation results obtained at \( \gamma = 10 \), the reference curve is the BER at \( \gamma = 0 \). All equalizers have 10\( M_s + 1 \) taps: 5\( M_s \) causal taps, 5\( M_s \) anti-causal taps and one main tap. The decision feedback equalizer has 5 taps, the non-linear decision feedback supplements this with 10 taps of non-linear symbol combinations. We have to remark that, to be able to simulate the BER semi-analytically, the feedback in the decision feedback equalizers is performed with the actual symbols instead of the estimated symbols. The actual performance of the decision feedback equalizers will hence be slightly worse because of error propagation.

We conclude that the decision feedback equalizer clearly outperforms the linear equalizer and that adding non-linear feedback improves the performance even further. It is also clear that increasing \( M_s \) while keeping the equalizer type fixed, increases the performance. Note that going from the linear equalizer to a decision feedback equalizer with \( M_s = 1 \) corresponds to a larger performance gain than going to a fractionally spaced equalizer with \( M_s = 4 \). The former change corresponds to an increase of only 5 equalizer taps whereas the latter increases the total number of taps with 30!

IV. CONTINUOUS TIME MMSE EQUALIZER

A. Algorithm

The MMSE criterion can also be used to design continuous time equalizers. We study equalizers consisting of a tapped delay line and programmable tap amplifiers. As the delay line and tap amplifiers are analog, we have to take the tap frequency responses into account. Neglecting the differences between the tap frequency responses and other non-idealities, we can formulate the following practical 6 step algorithm:
Step 1 Apply a known symbol sequence $a$ to the system, set one tap of the equalizer to its maximal value and the others to zero. Sample the signal after the equalizer to get $q(t)$.

Step 2 Fill $R$ and $r$ based on $q(t)$, $a$ and the current value of the shift $\tau$ between $q(t)$ and $a$.

Step 3 Calculate the coefficients $h$ corresponding to the system of equations $h = rR^{-1}$.

Step 4 Scale, adjust and quantize $h$ to the programmable values $h_{\text{programmable}}$.

Step 5 Readjust and scale $h_{\text{programmable}}$ back to $h_{\text{quantized}}$ to calculate the corresponding MSE.

Step 6 Carry out steps 2-5 for a range of $\tau$-values and select the result $(\tau, h_{\text{programmable}})$ that yields the lowest MSE. In this step it is advisable to smooth the MSE as a function of $\tau$ to get consistent results.

B. Measurement Results

We test our algorithm on a PON emulation setup and a Hittite HMC6545LP5E [1] continuous-time equalizer with 9 programmable, 18 ps-spaced taps. In Fig. 3 the power gain obtained by equalizing the signal, i.e. the decrease in optical power needed to get the same BER as before equalization, is shown for different settings of the dispersion $DL$.

![Fig. 3. Power gain Hittite HMC6545LP5E equalizer and MMSE coefficients.](image)

We see that the gain of the equalizer depends on two factors: the dispersion and the BER. The former is of course because more gain can be achieved if dispersion deteriorated the signal more. The latter is because the distance between the equalized and non-equalized BER curves increases for decreasing BER. This we also see in Fig.2, especially when a BER floor is present.

V. Conclusion

Several strategies for discrete time MMSE equalization of the PON communication signal were investigated. A criterion for the existence of a ZF equalizer for non-linear channels was derived. A practical algorithm for continuous time equalizers was formulated. Simulations were performed to investigate the discrete time equalizer performance, measurements verify the performance of the continuous time equalizer.
## Contents

1 Introduction .................................................. 1
   1.1 Overview ................................................ 1
   1.2 PON .................................................... 2
   1.3 Physical background .................................... 3

2 System model ............................................... 6
   2.1 Model .................................................. 6
      2.1.1 System without dispersion ......................... 8
      2.1.2 Influence of dispersion ............................ 11
   2.2 Details simulation ...................................... 17
      2.2.1 Implementation $p_{fib}$ .......................... 17
      2.2.2 Simulation of the BER ............................ 18

3 Discrete MMSE equalizer .................................. 20
   3.1 Linear equalizer ........................................ 20
      3.1.1 Algorithm ......................................... 21
      3.1.2 Simulation results ................................. 23
      3.1.3 Comparison with linear system .................... 29
   3.2 Fractionally spaced linear equalizer .................. 31
      3.2.1 Concept ........................................... 31
      3.2.2 Algorithm ......................................... 33
      3.2.3 Simple example .................................... 34
      3.2.4 Simulation results ................................. 42
   3.3 Decision feedback equalizer ............................ 47
      3.3.1 Concept ........................................... 47
      3.3.2 Algorithm ......................................... 48
      3.3.3 Simulation results ................................. 48
## CONTENTS

3.4 Non-linear decision feedback ........................................ 56
   3.4.1 Concept ....................................................... 56
   3.4.2 Algorithm .................................................... 56
   3.4.3 Simulation results ............................................ 57
3.5 Disadvantages of MSE criterion ..................................... 58

4 Continuous time equalizer ........................................... 61
   4.1 Test setup ....................................................... 61
      4.1.1 Transmitter ................................................. 62
      4.1.2 Channel ..................................................... 62
      4.1.3 Receiver .................................................... 63
      4.1.4 Equalizer .................................................... 63
   4.2 Algorithm ....................................................... 65
   4.3 Results .......................................................... 69
      4.3.1 BER ........................................................ 69
      4.3.2 Eye diagrams ............................................... 76
      4.3.3 Equalizer responses ....................................... 76

5 Conclusions .......................................................... 78
   5.1 Conclusions .................................................... 78
   5.2 Future work ..................................................... 81

Appendices ............................................................. 85

A Fiber impulse response ............................................. 86
   A.1 Calculation $h_{fib}$ .......................................... 86
   A.2 Gaussian integral with complex offset ......................... 87

B Derivation expression for $s(t)$ ................................... 88

C Calculation of equalizer coefficients ................................ 89
   C.1 Calculation of expression MSE ................................ 89
   C.2 Linear equalization with multiple samples per symbol .... 90
   C.3 Linear equalization with decision feedback .................. 90

D Measurement results ................................................ 93
   D.1 Eye diagrams per dispersion setting ......................... 93
   D.2 BER plots per dispersion setting ............................ 96
List of Figures

1.1 Architecture of a typical PON network. ............................................. 3

1.2 Typical optical fiber. ................................................................. 4

2.1 Block diagram simulated system. ................................................ 6

2.2 Magnitude optimized $H_{TIA}$. .................................................. 10

2.3 Optimized $h_{tot}(t)$ and samples at $\{\tau + mT\}$. ......................... 11

2.4 Real and imaginary part of $h_{fib}(t)$ for $\gamma = 3$. ....................... 12

2.5 Pulses $h_{NRZ}(t)$ and $p_{fib}(t)$ for $\gamma = 3$. ............................. 13

2.6 Longer segment of $p_{fib}(t)$ for $\gamma = 3$. .................................. 13

2.7 Eye diagrams at scaled TIA output. .............................................. 14

2.8 Eye opening and power penalty in function of $\gamma$. ....................... 15

2.9 Simulated bit error rate (BER) in function of $\gamma$. ....................... 16

2.10 Schematic of $s(t)$. ............................................................... 16

3.1 BER curves for different $\gamma$. ................................................ 21

3.2 Diagram of system with linear equalizer. ..................................... 21

3.3 Eye diagram and MSE for $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$. .......... 24

3.4 MSE as function of the sample moment for different number of filter taps for $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$. .............................. 26

3.5 BER curves for $\gamma = 1$. ......................................................... 27

3.6 BER curves for $\gamma = 4$. ......................................................... 28

3.7 BER curves for $\gamma = 6$ and $\gamma = 7$. ..................................... 29

3.8 Influence non-linearity on equalizer performance. .......................... 30

3.9 Influence filter length on linear system. ..................................... 30

3.10 Diagram of system with fractional linear equalizer. ....................... 31

3.11 Linear system with fractional equalizer. .................................. 32

3.12 Non-linear system with fractional equalizer. ............................. 33
3.13 Diagram of simple system. ............................................. 34
3.14 $p_{alt}(t)$ for $\gamma = 2$, $\mu = \ln(1/2)$ and $\beta = 1$. ................................. 35
3.15 Schematic representation of the error. ............................................. 36
3.16 BER for $\gamma = 2$ and $\mu = \ln(5)$. ............................................. 37
3.17 BER for $\gamma = 2$ and $\mu = \ln(2)$. ............................................. 39
3.18 Histogram of the unequalized system for $\gamma = 2$ and $\mu = \ln(5/3)$. .................. 40
3.19 Histogram of the equalized system with one tap for $\gamma = 2$ and $\mu = \ln(5/3)$. ....... 40
3.20 Histogram of the equalized system with $M_s = 2$ for $\gamma = 2$ and $\mu = \ln(5/3)$. .... 41
3.21 Histogram of the equalized system with $M_s = 4$ for $\gamma = 2$ and $\mu = \ln(5/3)$. .... 41
3.22 Histogram of the equalized system with $M_s = 8$ for $\gamma = 2$ and $\mu = \ln(5/3)$. .... 42
3.23 BER for $\gamma = 2$, $\mu = \ln(5/3)$ and different $M_s$. ............................................. 42
3.24 Histogram of the error for $M_s = 1$ (left), 2 (center) and 4 (right) for $\gamma = 1$. ....... 44
3.25 Histograms of the errors for $M_s = 1$ (left), 2 (center) and 4 (right) for $\gamma = 4$. .... 44
3.26 Histograms of the errors for $M_s = 1$ (left), 2 (center) and 4 (right) for $\gamma = 7$. .... 45
3.27 Histogram of the error for $M_s = 32$ and $\gamma = 7$. ............................................. 45
3.28 BER as function of $E_b/N_0$ with $\gamma = 1$ and $M_s = 1$, 2, 4, 8 and 16. ............... 45
3.29 BER as function of $E_b/N_0$ with $\gamma = 4$ and $M_s = 1$, 2, 4, 8 and 16. ............... 46
3.30 BER as function of $E_b/N_0$ with $\gamma = 7$ and $M_s = 1$, 2 and 4 .......................... 46
3.31 BER as function of $E_b/N_0$ with $\gamma = 7$ and $M_s = 4$ and 32 .......................... 47
3.32 Diagram of system with linear equalization and decision feedback. ......................... 48
3.33 Histograms of the error for $h_{FB}$ with 0, 1 and 5 taps for $\gamma = 1$. ..................... 49
3.34 Histograms of the error for $h_{FB}$ with 0, 1 and 5 taps for $\gamma = 4$. ..................... 50
3.35 Histograms of the error for $h_{FB}$ with 0, 1 and 5 taps for $\gamma = 7$. ..................... 50
3.36 BER curve for $\gamma = 1$ ............................................. 50
3.37 BER curve for $\gamma = 4$ for $M_s = 1$ (top figure) and for $M_s = 4$ (bottom figure) .... 51
3.38 BER curve for $\gamma = 7$ ............................................. 52
3.39 BER curve for $\gamma = 8$ ............................................. 52
3.40 BER curve for $\gamma = 10$ ............................................. 53
3.41 BER for $\gamma = 4$ and decision feedback with $\hat{a}$. ............................................. 54
3.42 BER for $\gamma = 7$ and decision feedback with $\hat{a}$. ............................................. 54
3.43 Length of error events for $\gamma = 4$ ............................................. 55
3.44 Length of error events for $\gamma = 7$ ............................................. 55
3.45 Visualization of non-linear terms in the feedback ............................................. 57
3.46 BER curves in case of non-linear feedback and $\gamma = 10$. ............................................. 57
3.47 BER curves in case of non-linear feedback and $\gamma = 15$. .......................... 58
3.48 Histograms magnitude error unequalized system with $p_{alt}(t)$. ......................... 59
3.49 Histograms magnitude error equalized system with $p_{alt}(t)$. .......................... 59
3.50 BER curves of system with $p_{alt}(t)$. .............................................................. 60
4.1 Test setup. ............................................................................................................ 62
4.2 Frequency response per tap, measured at coefficient value 63. ............................... 64
4.3 Fitted relationship between coefficient and small signal gain for every tap. ............. 65
4.4 Comparison mean square error (MSE): measurement number 1 and 2, resulting coefficients: ................................................................................................................. 68
4.5 Comparison smoothed MSE: measurement number 1 and 2, resulting coefficients: 68
4.6 BER of reference system. ......................................................................................... 69
4.7 BER of equalized system. ......................................................................................... 70
4.8 BER of system equalized with coefficients of system 500 ps/nm. ......................... 70
4.9 BER at dispersion of 300 ps/nm. ............................................................................. 71
4.10 Comparison MSE. ................................................................................................. 72
4.11 Schematic overview receiver. ................................................................................. 73
4.12 Power gain of equalized system. ........................................................................... 74
4.13 Power gain of system equalized with coefficients of system 500 ps/nm. .............. 75
4.14 Eye diagrams at a dispersion of 500 ps/nm. ......................................................... 76
4.15 Comparison equalizers at 100 ps/nm and 500 ps/nm. ............................................. 77
D.1 Eye diagrams at a dispersion of 0 ps/nm. .............................................................. 93
D.2 Eye diagrams at a dispersion of 100 ps/nm. ......................................................... 94
D.3 Eye diagrams at a dispersion of 200 ps/nm. ......................................................... 94
D.4 Eye diagrams at a dispersion of 300 ps/nm. ......................................................... 94
D.5 Eye diagrams at a dispersion of 400 ps/nm. ......................................................... 95
D.6 Eye diagrams at a dispersion of 500 ps/nm. ......................................................... 95
D.7 BER at dispersion of 0 ps/nm. ................................................................................. 96
D.8 BER at dispersion of 100 ps/nm. ............................................................................. 97
D.9 BER at dispersion of 200 ps/nm. ............................................................................. 97
D.10 BER at dispersion of 300 ps/nm. .......................................................................... 98
D.11 BER at dispersion of 400 ps/nm. .......................................................................... 98
D.12 BER at dispersion of 500 ps/nm. .......................................................................... 99
List of Tables

1.1 Overview of the most important PON standards. .......................... 2
2.1 Fiber lengths corresponding to a given bit rate and $\gamma$. ................. 7
3.1 Possible values of $r(k)$ for $\gamma = 2$. ........................................ 36
3.2 Possible values of $r(k)$ for $\gamma = 2$ and $A = -1/5$. ..................... 37
3.3 Possible values of $r(k)$ for $\gamma = 2$ and $A = -1/2$. ..................... 38
3.4 Possible values of $r(k)$ for $\gamma = 2$ and $A = -3/5$. ..................... 39
3.5 MSE in case for $\gamma = 1$ and multiple samples per symbol. ............... 43
3.6 MSE in case for $\gamma = 4$ and multiple samples per symbol. ............... 43
3.7 MSE in case for $\gamma = 7$ and multiple samples per symbol. ............... 43
3.8 MSE different configuration with decision feedback. ......................... 49
3.9 Feedback filter coefficients at $E_b/N_0 = 15$ dB. .......................... 54
Acronyms

10G-EPON  10 Gbit/s ethernet PON
AC   alternating current
ADC  analog to digital converter
AGC  automatic gain control
BER  bit error rate
BPON broadband PON
CO   central office
CW   continuous wave
EDFA erbium doped fiber amplifier
EPON ethernet PON
FBT  fused biconical taper
FFT  fast fourier transform
FTTH fiber to the home
GPON gigabit-capable PON
IFFT inverse fast fourier transform
ISI  intersymbol interference
MAP  maximum a posteriori probability
ML   maximum-likelihood
MMSE minimal mean square error
MSE  mean square error
MZM  Mach-Zehnder modulator
NRZ  non-return-to-zero
Acronyms

OLT optical line termination
ONU optical network units
PD photodetector
PLC planar lightwave circuit
PMD polarization mode dispersion
PON passive optical network
PRBS pseudorandom binary sequence
PSD power spectral density
RZ return-to-zero
SNR signal to noise ratio
TDM time division multiplexing
TDMA time division multiple access
TIA transimpedance amplifier
XG-PON next generation PON
ZF zero forcing
List of Symbols

$A$ gain for $a(k-1)$ and $a(k+1)$ in simple example
$a$ data symbol
$\hat{a}$ estimated data symbol
$\alpha$ fiber loss in dB/km
$c_{off}$ offset of forward filter
$D$ chromatic dispersion parameter
$d_{eye}$ vertical eye opening
$DL$ dispersion parameter of the dispersion module
$\zeta$ damping factor
$d$ decision threshold
$e$ error signal
$E_b$ Energy per bit at output of the fiber
$e_{in}$ optical field at input of the fiber
$e_{out}$ optical field at output of the fiber
$\epsilon$ relative accuracy
$f_s$ sample frequency of the simulation
$\gamma$ dispersion parameter
$h_{FB}$ feedback equalizer filter
$h_{FF}$ forward equalizer filter
$h_{fib}$ impulse response of the fiber
$h_{NRZ}$ NRZ pulse
$h_{tap}$ impulse response of one tap of continuous-time equalizer
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i_n$</td>
<td>noise current</td>
</tr>
<tr>
<td>$i_{pd}$</td>
<td>current after photodetector</td>
</tr>
<tr>
<td>$i_{pd,n}$</td>
<td>current at input transimpedance amplifier</td>
</tr>
<tr>
<td>$L$</td>
<td>fiber length</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>optical wavelength</td>
</tr>
<tr>
<td>$M$</td>
<td>length $p_{fib}$ in simulation</td>
</tr>
<tr>
<td>$M_s$</td>
<td>number of samples per symbol period</td>
</tr>
<tr>
<td>$\mu$</td>
<td>decay parameter in simple example</td>
</tr>
<tr>
<td>$n$</td>
<td>refractive index</td>
</tr>
<tr>
<td>$N_0$</td>
<td>noise power</td>
</tr>
<tr>
<td>$N_{FF}$</td>
<td>length of forward filter</td>
</tr>
<tr>
<td>$N_s$</td>
<td>samples per symbol in simulation</td>
</tr>
<tr>
<td>$\omega_0$</td>
<td>natural pulsation</td>
</tr>
<tr>
<td>$p_{fib}$</td>
<td>NRZ pulse applied to fiber</td>
</tr>
<tr>
<td>$P_{PD}$</td>
<td>measured power at the input of photodiode</td>
</tr>
<tr>
<td>$r$</td>
<td>received voltage after transimpedance amplifier</td>
</tr>
<tr>
<td>$s$</td>
<td>signal component of $r$</td>
</tr>
<tr>
<td>$\sigma_0^2$</td>
<td>variance</td>
</tr>
<tr>
<td>$T$</td>
<td>symbol period</td>
</tr>
<tr>
<td>$\tau$</td>
<td>sampling delay</td>
</tr>
<tr>
<td>$T_{tap}$</td>
<td>tap spacing continuous-time equalizer</td>
</tr>
<tr>
<td>$u$</td>
<td>output linear forward equalizer</td>
</tr>
<tr>
<td>$w$</td>
<td>window in simple example</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Overview

We begin this thesis with an overview of the context of the system we investigate: passive optical networks (PONs) are introduced and we present a brief physical background on optical fibers. In the next chapter a simplified model of a PON communication system is developed where the most important components are the fiber chromatic dispersion and a non-linearity in the photodetector (PD). The model is used to optimize some system parameters and serves as a basis for our computer simulations. Due to the dispersion and PD, the performance of the system is deteriorated.

To mitigate the performance loss, we investigate discrete time equalization of the electrical received signal. Linear equalization, linear fractionally spaced equalization and (non-linear) decision-feedback equalization based on the minimal mean square error (MMSE) criterion are treated. Their performance, especially in the case of a bit error rate (BER) floor corresponding to a closed eye-diagram, is studied and compared. We end this chapter with a short discussion of the disadvantages of the MMSE criterion.

The next chapter is devoted to continuous time equalization. A PON emulation test setup with a continuous time equalizer is described and discussed. We adapt the discrete equalizer algorithms to a practical algorithm for the continuous time equalizer, taking its non-idealities into account. Using our algorithm, we equalize the PON emulation test setup and measure the performance improvement we can achieve.
In a last chapter we give a short overview of what we achieved and discuss the possibilities for future work.

### 1.2 PON

To accommodate the ever-increasing bandwidth demands, network operators have to move the optical-electrical conversion ever closer to their subscribers. Deployment of PONs to offer fiber to the home (FTTH) connections is an important step in this evolution. Figure 1.1 shows the typical architecture of a PON with an optical line termination (OLT) at the network operator’s central office (CO), a passive optical splitter and a number of optical network units (ONUs) at the premises of the end users. The OLT and ONU serve as interfaces between the PON and, respectively, the core network/the end user’s devices. In between we have the actual fiber and the optical splitter, the fiber is a single mode fiber and the splitter generally has a ratio of 1:32 but other ratios are also possible, provided the used PON standard supports them. The point-to-multipoint topology minimizes the number of optical transceivers and the passivity of the intermediate network nodes (e.g. the optical splitter) means we do not need to provide power alongside the fiber, which makes it low cost, easy to maintain and reliable. These are two big advantages that explain the success of PONs.

Multiple PON standards exist, the most important ones are summarized in table 1.1 [1–5].

<table>
<thead>
<tr>
<th>Standard</th>
<th>Typical bit rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Downstream</td>
</tr>
<tr>
<td>ITU-T G.983 broadband PON (BPON)</td>
<td>622 Mbit/s</td>
</tr>
<tr>
<td>ITU-T G.984 gigabit-capable PON (GPON)</td>
<td>2.488 Gbit/s</td>
</tr>
<tr>
<td>ITU-T G.987 next generation PON (XG-PON)</td>
<td>10 Gbit/s</td>
</tr>
<tr>
<td>IEEE 802.3ah ethernet PON (EPON)</td>
<td>1 Gbit/s</td>
</tr>
<tr>
<td>IEEE 802.3av 10 Gbit/s ethernet PON (10G-EPON)</td>
<td>10 Gbit/s</td>
</tr>
</tbody>
</table>

All these standards use wavelength-division duplexing to offer both upstream and downstream over the same fiber. BPON, GPON and EPON use the same wavelengths: 1490 nm for downstream traffic, 1310 nm for upstream traffic and 1550 nm for optional overlay services (e.g. video). To avoid interference between the users in the upstream time division multiple access
(TDMA) is generally used, where every subscriber must coordinate with the CO when it can send its information. In the downstream the CO uses a tag-based method to broadcast all the information to all the subscribers. Each subscriber then just selects the data corresponding to its address tag. This is an implementation of time division multiplexing (TDM) [6].

![Architecture of a typical PON network.](image)

The standards prescribe a certain optical budget, based on the maximum allowable BER. This loss budget limits the amount of users and/or the maximal OLT-ONU distance. The GPON standard, for example, prescribes 15-30 dB attenuation in class C. If we want to serve 32 subscribers (15 dB + max 4 dB excess loss) and assume $\alpha = 0.5$ dB/km fiber loss (ANSI/EIA/TIA 568-C [7]), this corresponds to a maximum distance of approximately 20 km. Increasing this optical budget is possible in many ways: applying extra coding, better optics (e.g. fiber with lower losses), etc. In this thesis we will study electrical equalization of the received signal to decrease the BER for a given received optical power. As our equalization is performed at the physical layer, we are not interested in whatever happens to the signal after equalization. Possible coding of the signal is ignored, we only try to minimize the BER after equalization. Note that this makes advanced equalization (e.g. Turbo equalization [8]) impossible.

### 1.3 Physical background

In this section we briefly describe the optical fiber and the dispersion phenomena based on [9,10].

All optical fibers possess a cylindrical geometry with a core (refractive index $n_1$) having a (slightly) larger refractive index than the cladding material (refractive index $n_2$), see Figure 1.2. In this structure it is possible for light to propagate through the internal structure by means of total internal reflection. Depending on the size of the fiber and the difference in refraction index, there are different possible modes for the light to propagate.
1.3 Physical background

Figure 1.2: Typical optical fiber.

Although the bandwidth of the optical fiber is typically very wide (> 10 THz), the bandwidth of the modulated signal will be limited due to the dispersion. Dispersion means that the phase velocity of the wave traveling in the structure depends on its frequency. Hence the propagation delay in the time domain will be dependent on the wavelength, resulting in distortion of the transmit pulse.

In the optical fiber there are different causes for dispersion. First of all we will suffer from dispersion when we have multiple propagating modes in the fiber. Each mode will have its own traveling path and its corresponding propagation delay, which will lead to spreading of the pulse. The solution to avoid this distortion is using a single-mode fiber. This is a fiber with only one propagation mode and hence no modal dispersion. In this thesis we assume we have a single-mode fiber, so we do not model this dispersion.

A second type of dispersion is called chromatic dispersion. The cause here is that the group velocity depends on the wavelength, meaning that if the source has a certain spectral width due to some modulation, the different wavelength components will travel at different speeds and the pulse broadens. In practice, this chromatic dispersion is specified by the change in group delay per nm wavelength and km length:

$$D = \frac{1}{L} \frac{\partial \tau}{\partial \lambda}$$  \hspace{1cm} (1.1)

where $D$ is the chromatic dispersion parameter, $L$ the fiber length, $\tau$ the group delay and $\lambda$ the wavelength.

A third dispersion type, polarization mode dispersion (PMD) is the result of polarization modes propagating at different speeds. In a single-mode fiber this is caused by the two orthogonal polarizations of the fundamental mode. Ideally the optical fiber core is perfectly circular and thus has the same refractive index for both polarization modes. In a real fiber, however, mechanical and thermal stresses introduced during manufacturing result in asymmetries in the fiber core geometry, which causes a different refractive index for each polarization mode. In this thesis we do not model PMD but conceptually the equalizers we design can also deal with this type of dispersion.
Finally we mention that the optical fiber also will suffer from losses. The attenuation of the optical signal is a result of scattering, absorption and other effects. As these phenomena are wavelength dependent, the loss is not the same for all wavelengths. One of the main reason to use the wavelengths of 1300 nm and 1550 nm in optical communications, is that silica glass shows a minimum loss at those wavelengths. In this thesis we do not directly model the fiber loss, since we are only interested in the possibility to equalize the waveforms, independent of the signal level.
Chapter 2

System model

2.1 Model

To evaluate and characterize the equalizers we design, we study and simulate a model of the PON communication system and the equalization algorithms. We only include the essential components of the PON in our model, a schematic overview is shown in Figure 2.1.

At the transmitter we start with a stream of data bits $a_m \in \{0, 1\}$ with a mean value of $E[a_m] = 0.5$ and a rate of $1/T$. We do not have any information about the possible coding of this data, we assume that the ones and zeros are independent and equally probable. These data bits are mapped to a unipolar non-return-to-zero (NRZ) signal which modulates an optical carrier generated by a continuous wave (CW) laser with wavelength $\lambda$ through a Mach-Zehnder modulator (MZM). The resulting envelope $e_{in}(t)$ of the optical field at the MZM output is given by

$$e_{in}(t) = \sum_m a_m h_{NRZ}(t - mT)$$

(2.1)

where the NRZ pulse $h_{NRZ}(t)$ equals 1 for $0 \leq t \leq T$. In the simulation we do not take the passive optical splitter into account because we ignore any reflection so its only contribution is a fixed attenuation. We can do this because both fused biconical taper (FBT) and planar
lightwave circuit (PLC) splitters generally have return losses and directivity $\geq 50$ dB (e.g. [11] and [12]). The single mode fiber introduces chromatic dispersion, the envelope $e_{\text{out}}(t)$ at the output of the fiber is represented as

$$e_{\text{out}}(t) = \int h_{\text{fib}}(u)e_{\text{in}}(t-u)du$$

where $h_{\text{fib}}(u)$ is the inverse Fourier transform of the fiber transfer function $H_{\text{fib}}(f) = \exp(-j\gamma f^2 T^2)$. The relation between $\gamma$ and $\lambda$, $L$ and $T$ is given by $\gamma = \pi D(\lambda) \frac{\lambda^2}{cT^2} L$ with $D(\lambda)$ the chromatic dispersion parameter. This parameter depends on both the wavelength $\lambda$ and the design of the fiber (ITU-T standards G.652 up to G.655 [13–16] describe different types of single mode fiber), we assume that $D = 15$ ps/(km.nm). Table 2.1 shows the bit rates and fiber lengths corresponding to the $\gamma$ we will investigate.

<table>
<thead>
<tr>
<th>Bit Rate</th>
<th>$\gamma = 1$</th>
<th>$\gamma = 3$</th>
<th>$\gamma = 6$</th>
<th>$\gamma = 7$</th>
<th>$\gamma = 8$</th>
<th>$\gamma = 10$</th>
<th>$\gamma = 15$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 Gbit/s</td>
<td>26.5 km</td>
<td>79.5 km</td>
<td>159.0 km</td>
<td>185.5 km</td>
<td>212.0 km</td>
<td>265.0 km</td>
<td>397.5 km</td>
</tr>
<tr>
<td>20 Gbit/s</td>
<td>6.6 km</td>
<td>19.9 km</td>
<td>39.7 km</td>
<td>46.4 km</td>
<td>53.0 km</td>
<td>66.2 km</td>
<td>99.4 km</td>
</tr>
<tr>
<td>50 Gbit/s</td>
<td>1.1 km</td>
<td>3.2 km</td>
<td>6.4 km</td>
<td>7.4 km</td>
<td>8.5 km</td>
<td>10.6 km</td>
<td>15.9 km</td>
</tr>
<tr>
<td>100 Gbit/s</td>
<td>0.3 km</td>
<td>0.8 km</td>
<td>1.6 km</td>
<td>1.9 km</td>
<td>2.1 km</td>
<td>2.6 km</td>
<td>4.0 km</td>
</tr>
</tbody>
</table>

The modulated light at the output of the fiber is converted into an electrical signal by means of a PD, which outputs a current $i_{PD}(t)$ given by $|e_{\text{out}}(t)|^2$. Note that the system is non-linear due to this PD. To the PD current we add a zero-mean stationary white Gaussian noise current $i_{n}(t)$ with power spectral density (PSD) $N_0/2$. Next, we apply $i_{PD,n}(t) = i_{PD}(t) + i_{n}(t)$ to a unity gain transimpedance amplifier (TIA) with second-order frequency response

$$H_{TIA}(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0s + \omega_0^2}$$

where $\omega_0$ is the natural pulsation and $\zeta$ is the the damping factor. The voltage at the output of the TIA is given by

$$v_{TIA}(t) = \int h_{TIA}(u)i(t-u)du$$

with

$$h_{TIA}(u) = \frac{\omega_0}{\sqrt{1-\zeta^2}}\exp(-\zeta\omega_0 u)\sin(\sqrt{1-\zeta^2}\omega_0 u)$$
the impulse response of a second-order low-pass system. Sampling of \( v_{TIA}(t) \) yields \( r(mM_s+k) = v_{TIA}(mT+kT/M_s + \tau) \) with \( M_s \) the number of samples per symbol period and \( \tau \) an appropriate sampling delay. These samples are the input to our equalizer. The output of the equalizer is decimated and symbol-by-symbol detection is performed afterwards by comparing the decimated equalizer samples to the decision threshold \( d = 0.5 \).

Note that all continuous-time signals are simulated in discrete time with a sample rate of \( f_s = N_s/T \) with \( N_s = 64 \) for all simulations.

Because we have kept our model as straightforward as possible, our simulation has only a limited number of parameters other than the equalizer parameters: \( \gamma, T, \omega_0, \zeta \) and \( E_b/N_0 \). The last parameter conveniently lumps the CW laser power, MZM losses, splitter attenuation, fiber attenuation, PD responsivity and TIA noise. \( E_b \) is the energy per bit at the output of the fiber scaled by the PD responsivity and \( N_0/2 \) is the power spectral density of the white noise added at the input of the TIA. We will optimize the TIA parameters \( \omega_0 \) and \( \zeta \) for a minimum BER in the system without dispersion and will keep \( T \) fixed at 10 GHz. As a result we only have the two parameters \( \gamma \) and \( E_b/N_0 \) left that depend on the bit rate \( 1/T \), the fiber length \( L \), the number of subscribers \( N_{\text{users}} \) and the fiber loss \( \alpha \) as follows:

\[
\gamma = \gamma_0\frac{T^2}{T_0^2} \frac{L}{L_0} \tag{2.6}
\]

\[
\frac{E_b}{N_0} = \left( \frac{E_b}{N_0} \right)_0 - 10 \log_{10} \frac{N_{\text{users}}}{N_{\text{users},0}} - \alpha(L-L_0) \tag{2.7}
\]

Increasing the bit rate \( 1/T_0 \) with a factor \( \sqrt{10} \) corresponds to increasing \( \gamma \) with a factor 10, doubling the amount of subscribers corresponds to decreasing \( E_b/N_0 \) by 3 dB and increasing the fiber length from \( L_0 \) to \( L \) corresponds to increasing \( \gamma \) with a factor \( L/L_0 \) and decreasing \( E_b/N_0 \) by \( \alpha(L-L_0) \). Using equations 2.6 and 2.7 we can translate simulation results in function of \( \gamma \) and \( E_b/N_0 \) to results in function of \( T, L \) and \( N_{\text{users}} \).

### 2.1.1 System without dispersion

First we consider the system without dispersion (\( \gamma = 0 \)) and \( M_s = 1 \). The system reduces to a linear system because the unipolar NRZ signal (\( \in \{0,1\} \)) is not affected by the PD non-linearity and we get

\[
i_{PD}(t) = \sum_m a_m h_{NRZ}(t - mT) \tag{2.8}
\]
The sample \( r_k = r(k) \) can be decomposed into contributions from \( i_{PD}(t) \) and \( i_N(t) \) as \( r_k = s_k + n_k \). As \( i_N(t) \) is zero-mean stationary white Gaussian noise with PSD \( N_0/2 \) it follows that the noise contribution \( n_k \) is zero-mean Gaussian with variance \( \sigma_n^2 \) given by
\[
\sigma_n^2 = \frac{N_0}{2} \int_{-\infty}^{\infty} |H(j2\pi f)|^2 df
= \frac{N_0 \omega_0}{2} \frac{\omega}{4\zeta}
\]
with \( \omega/4\zeta \) the noise bandwidth of the second order low-pass filter. The signal contribution is given by
\[
s_k = \sum_m a_m h_{tot}(kT + \tau - mT)
= 0.5 + \sum_m \Delta_k - m h_{tot}(mT + \tau)
\]
with \( h_{tot}(t) = h_{step}(t) - h_{step}(t - T) \) the response of second order low pass filter to \( h_{NRZ}(t) \) and \( \Delta_m = a_m - 0.5 \in \{-0.5, 0.5\} \). The step response \( h_{step}(t) \) of the filter is given by
\[
h_{step}(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 - \frac{\exp(-\zeta\omega_0 t)}{\sqrt{1 - \zeta^2}} \sin(\sqrt{1 - \zeta^2} \omega_0 t) & \text{if } t \geq 0 \end{cases}
\]
We can decompose \( s_k \) into an offset term, a useful term and an intersymbol interference (ISI) term as follows:
\[
s_k = 0.5 + \underbrace{\Delta_k h_{tot}(\tau)}_{\text{useful term}} + \sum_{m \neq 0} \underbrace{\Delta_k - m h_{tot}(mT + \tau)}_{\text{ISI}_k(\tau)}
\]
The signal values \( s_k \) are located in the intervals \([s_{min,+}, s_{max,+}]\) or \([s_{min,-}, s_{max,-}]\) when \( \Delta_k = 0.5 \text{sgn}(h_{tot}(\tau)) \) or \( \Delta_k = -0.5 \text{sgn}(h_{tot}(\tau)) \) respectively with
\[
s_{min,+} = 0.5 + 0.5|h_{tot}(\tau)| - \text{ISI}_{max}(\tau)
\]
\[
s_{max,+} = 0.5 + 0.5|h_{tot}(\tau)| + \text{ISI}_{max}(\tau)
\]
\[
s_{min,-} = 0.5 - 0.5|h_{tot}(\tau)| - \text{ISI}_{max}(\tau)
\]
\[
s_{max,-} = 0.5 - 0.5|h_{tot}(\tau)| + \text{ISI}_{max}(\tau)
\]
and \( \text{ISI}_{max}(\tau) = 0.5 \sum_{m \neq 0} |h_{tot}(mT + \tau)| \) the maximum ISI magnitude. We can see that the intervals are disjoint when the magnitude of the useful term, \( 0.5|h_{tot}(\tau)| \), is larger than the maximum ISI magnitude \( \text{ISI}_{max}(\tau) \). Detection then comes down to comparing \( r_k \) to a threshold value of 0.5:
\[
\hat{\Delta}_k = \begin{cases} \text{sgn}(h_{tot}(\tau)) & \text{if } r_k > 0.5 \\ -\text{sgn}(h_{tot}(\tau)) & \text{if } r_k \leq 0.5 \end{cases}
\]
From $ISI_{\text{max}}(\tau)$, equation 2.9 and equation 2.12 we can now derive an analytical expression for the worst-case BER

$$BER_{\text{max}} = Q\left(\frac{0.5|h_{\text{tot}}(\tau)| - ISI_{\text{max}}(\tau)}{\sigma_0}\right)$$

(2.14)

with $Q(x)$ the tail probability of the standard normal distribution and $0.5|h_{\text{tot}}(\tau)| - ISI_{\text{max}}(\tau)$ half the vertical eye opening at time instant $kT + \tau$ at the TIA output without noise.

This maximal BER is a function of $\tau$ and the TIA parameters $\omega_0$ and $\zeta$. From the joint optimization of these we find $\omega_{0,\text{opt}} = 3.1652/T$ and $\zeta_{\text{opt}} = 0.6707$ at $\tau = 66T/N_s = 33/32T$. In Figure 2.2 we see that the 3 dB cutoff frequency of our optimized TIA is located at $\approx 1/2T$. Figure 2.3 shows the resulting $h_{\text{tot}}(t)$ and its samples at $\{\tau + mT\}$. We get $h_{\text{tot}}(\tau) = 1.011$ and $ISI_{\text{max}}(\tau) = 0.018$ so half the vertical eye opening equals 0.99.

![Figure 2.2: Magnitude optimized $H_{TIA}$](image)

It is interesting to compare our second order filter to the optimum receive filter matched to the NRZ-pulse. In this case $h_{\text{tot}}$ is a triangular pulse of length $2T$

$$h_{\text{tot}}(t) = \begin{cases} 
0 & \text{if } t \leq 0 \text{ and } t \geq 2T \\
\frac{t}{T} & \text{if } 0 < t < T \\
2 - \frac{t}{T} & \text{if } T < t < 2T 
\end{cases}$$

(2.15)

and sampling at the optimal $\tau = T$ yields no ISI. The noise bandwidth of the matched filter is $1/T$ which is $\omega_0 T/4\zeta = 1.18$ times lower than that of the optimized second order filter. As a result the electrical power penalty of the second order filter is given by the penalty of the increased noise bandwidth (0.72 dB) and the penalty of the smaller eye opening (0.993 instead of 1, i.e. 0.06 dB) or 0.78 dB combined. This corresponds to an optical power penalty of 0.39 dB.
2.1 Model

From now on we will keep the TIA parameters fixed to the optimal values we found, even though dispersion and equalization will most likely shift them a bit. The reason for this is that introducing dispersion also introduces the non-linearity making the above analytical expression for $\text{BER}_{\text{max}}$ invalid. As a result we would need to perform a joint optimization of $\omega_0$, $\zeta$, $\tau$ and the equalizer coefficients.

### 2.1.2 Influence of dispersion

Now we introduce dispersion to the system by setting $\gamma \neq 0$. Recall that the fiber transfer function is equal to $H_{\text{fib}}(f) = \exp(-j\gamma f^2 T^2)$ and that the envelope of the optical field at the output of the fiber equals $e_{\text{out}}(t) = \int h_{\text{fib}}(u) e_{\text{in}}(t - u) du$. To determine $h_{\text{fib}}(t)$ we calculate the inverse Fourier transform of $H_{\text{fib}}(f)$ in Appendix A. We get

$$ h_{\text{fib}}(t) = \frac{1}{T} \sqrt{\frac{-j\pi}{\gamma}} e^{\frac{j\pi^2 t^2}{\gamma T^2}} $$

$$ = \sqrt{\frac{2}{2T}} \sqrt{\frac{\pi}{\gamma}} \left( \cos \left( \frac{\pi^2 t^2}{\gamma T^2} \right) + \sin \left( \frac{\pi^2 t^2}{\gamma T^2} \right) \right) + $$

$$ j \sqrt{\frac{2}{2T}} \sqrt{\frac{\pi}{\gamma}} \left( \sin \left( \frac{\pi^2 t^2}{\gamma T^2} \right) - \cos \left( \frac{\pi^2 t^2}{\gamma T^2} \right) \right) $$

In Figure 2.4 we plot the real and imaginary part of $h_{\text{fib}}(t)$ for $\gamma = 3$. Note the symmetry around zero, this is because the argument $t$ in equation 2.16 is squared. It is clear that $h_{\text{fib}}(t)$ is an infinite impulse response that does not tend to zero for $|t| \to +\infty$ making it impossible to implement in our computer simulation.
Instead we look at $p_{fib}(t)$, the response of the fiber to an NRZ pulse. The convolution $h_{fib}(t) \ast h_{NRZ}(t)$ can be interpreted as sending $h_{fib}(t)$ through a filter with frequency response $T \text{sinc}(fT) = \frac{\sin(\pi fT)}{\pi fT}$ which attenuates the high frequency oscillations present in $h_{fib}(t)$ for large $t/T$. Indeed, as the instantaneous phase $\phi(t)$ of $h_{fib}(t)$ equals $\frac{\pi^2 t^2}{\gamma T^2}$, the instantaneous frequency equals $f(t) = \frac{1}{2\pi} \frac{\partial \phi(t)}{\partial t} = \frac{\pi t}{\gamma T}$ and increases with $t$. As a result, $p_{fib}(t)$ tends to zero for $|t| \to +\infty$.

Figures 2.5 and 2.6 show both the NRZ pulse and the response of the fiber to it. Note that the pulses have been shifted to let the integer values on the x-axis correspond to sample instants $T/2 + kT$. We also show $T \text{sinc}(fT)$, with $f(t) = \frac{\pi t}{\gamma T}$ the instantaneous frequency of $h_{fib}(t)$ and scaled with the $\sqrt{\frac{\pi}{\gamma}}$ factor from equation 2.16, because we expect to recognize the sinc shape of the NRZ pulse in $p_{fib}(t)$ as a result of the relation between the instantaneous frequency and $t$.

Figures 2.5 and 2.6 show that the envelope of $p_{fib}(t)$ for large $t$ indeed follows this sinc.

We can also find this sinc envelope for large $t$ by approximating $(t+u)^2$ as $t^2 + 2tu$ in the integral of the convolution $h_{fib} \ast h_{NRZ}$. We get

$$|p_{fib}(t)| = \frac{1}{T} \sqrt{\frac{\pi}{\gamma}} \left| \int_0^T \exp \left( j \frac{\pi^2(t-u)^2}{\gamma T^2} \right) du \right|$$

$$\approx \frac{1}{T} \sqrt{\frac{\pi}{\gamma}} \left| \int_0^T \exp \left( j \frac{\pi^2(t^2 - 2tu)}{\gamma T^2} \right) du \right| \quad \text{for} \quad t \gg T$$

$$= \sqrt{\frac{\pi}{\gamma}} \left| \text{sinc} \left( \frac{\pi t}{\gamma T} \right) \right|$$

(2.17)

It is clear that the broadening of the NRZ-pulse to $p_{fib}$ leads to ISI at the input of the PD.
After the PD non-linearity we get

\[ i_{PD}(t) = \left| \sum_m a_m p_{fib}(t - mT) \right|^2 \]  

(2.18)
and the resulting signal contribution $s_k$ to the sampled version of the scaled $v_{TIA}(t)$ is given by

$$s_k = 2 \int h_{TIA}(u) \left| \sum_{m} a_{k-m} p_{fib}(mT + \tau - u) \right|^2 du$$

(2.19)

Note that the phase information of $e_{out}$ is lost because of the PD non-linearity.

In contrast to the system without dispersion where $s_k$ is linear in the data symbols (see equation 2.10), we now see in equation 2.19 that $s_k$ has become a quadratic function in the data symbols. Determining the worst case BER analytically is much more difficult in the system with dispersion because we need to find the worst-case data sequences that cause the maximum ISI for $a_k = 1$ and for $a_k = 0$ at a given sampling instant $kT + \tau$. Instead we simulate the eye diagram at the scaled TIA output in absence of noise because half the vertical eye opening $d_{eye}(\tau)$ at every sampling instant corresponds to the worst-case BER: $\text{BER}_{\text{max}} = Q \left( \frac{d_{\text{eye}}(\tau)}{\sigma_0} \right)$. In 2.7 we plot the eye diagrams for $\gamma = 0$ and $\gamma = 3$, the vertical line shows the location of the maximum vertical eye opening. We see that dispersion, obviously, closes the eye: $d_{\text{eye}}$ reduces from 0.993 (note: equal to what we determined analytically!) to 0.667 which corresponds to an electrical power penalty of 3.46 dB or an optical power penalty of 1.73 dB. Figure 2.8 shows the eye opening and electrical power penalty (compared to the system with $\gamma = 0$) in function of $\gamma$, we see that the eye is closed for $\gamma \approx 6$ and the power penalty hence goes to infinity.

![Eye diagrams for $\gamma = 0$ and $\gamma = 3$.](image)

(a) $\gamma = 0$, $d_{\text{eye}} = 0.993$

(b) $\gamma = 3$, $d_{\text{eye}} = 0.667$

Figure 2.7: Eye diagrams at scaled TIA output.

Note that the envelope of $p_{fib}$ is approximately zero at some time instants (cf. Figures 2.5 and 2.6) and hence we expect some choices of $\gamma$ where the ISI at the PD is limited. As we found that $p_{fib}$ approximately follows $\text{sinc} \left( \frac{\pi l}{T} \right)$, we find that this is happens at $\gamma = \pi / l$ with $l \in \mathbb{N}_{>0}$ because then $\text{sinc} \left( \frac{\pi l}{T} \right)$ is zero for $t = kT$, $k \neq 0$. Figure 2.9 shows the simulated BER of our system where we replace the TIA filter by a low-pass filter with cutoff frequency $32/T$ (i.e. influence TIA
negligible) and set $E_b/N_0 = 40$ dB. As expected, we see dips at integer divisions of $\pi$ (the dip at $\pi/3$ is not very deep because the resolution of $\gamma$ is limited to $\pi/64$ to keep the simulation time acceptable). This is not visible on Figure 2.8 because there the TIA filter has an optimized cutoff frequency of $\approx 1/2T$. As a result only the lower frequency components of $p_{fib}$ are important and Figure 2.5 shows that these are not zero at sinc zero-crossings.

Figure 2.8: Eye opening and power penalty in function of $\gamma$.

We can also interpret the influence of the chromatic dispersion followed by the non-linearity differently. In Appendix B we derive an expression for $s(t)$, the data component of $r(t)$. We get

$$s(t) = \sum_m a(m) \sum_{k=-M}^M a(m+k) h_{tot,k}(t - mT)$$

with $p_{fib}(t) \approx 0$ for $t \not\in [-MT, MT]$, $h_{tot,k}(t) = \int h_{tia}(t - u)p_{fib,k}^{(2)}(u) du$ and $p_{fib,k}^{(2)}(t) = p_{fib}(t)p_{fib}^*(t - kT)$. Note that $h_{tot,k}(t) = h_{tot,-k}^*(t)$.

From equation 2.20 it follows that $s(t)$ can be seen as a superposition of $2M + 1$ transmit pulses $h_{tot,k}(t)$ convolved with $2M + 1$ data sequences $a(m)a(m + k)$, $k \in [-M, M]$. A schematic representation can be found in Figure 2.10. Of these $2M + 1$ components of $s(t)$, we are only interested in the one corresponding to $k = 0$. Because we have $a(m) \in \{0, 1\}$, $a(m)a(m) = a(m)$.
2.1 Model

Figure 2.9: Simulated BER in function of $\gamma$.

Figure 2.10: Schematic of $s(t)$. 
2.2 Details simulation

In this section we give some more details on our implementation of the simulations in MATLAB®. During the implementation, our most important obstacles were the realization of \( p_{fib} \) and the simulation of the BER. For the former we discuss two questions: how to obtain the samples of \( p_{fib}(t) \) and where to truncate the infinite pulse. For the latter we describe a Monte Carlo approach and a semi-analytical approach [17,18].

2.2.1 Implementation \( p_{fib} \)

There are several methods to obtain the pulse \( p_{fib}(t) \) for computer simulation. The first and most obvious method is of course just convolving \( h_{NRZ}(t) \) and \( h_{fib}(t) \) together. We have an analytical expression for both, hence we can obtain a sampled version. However, at the time we wrote our simulation code we had not yet found the analytical \( h_{fib} \) expression so this is not the approach we followed.

Instead of convolution in the time domain we performed the equivalent operation of multiplying the spectra in the frequency domain. In our simulation we represent the continuous signals by signals sampled at the simulation sample frequency \( f_s = 1/T_s \). Hence we know that the spectrum of \( p_{fib}(kT_s) \) is the periodic extension of \( P_{fib}(f) = H_{fib}(f)H_{NRZ}(f) \). Here the problem is that the frequency response of both pulse goes from \( -\infty \) to \( +\infty \). Luckily we know that \( H_{NRZ} \) goes to zero for large \( f \) and as a result we can approximate the periodic extension with a finite sum \( P_{fib}(f) = \sum_{n=-Q+1}^{Q} P_{fib}(f - n/T_s) \). Now we can calculate the inverse fast fourier transform (IFFT) of this signal and truncate the result to \( t/T_s \in [-M, M] \) to obtain \( p_{fib}(kT_s) \). A good choice of \( M \) is discussed next.

In order to simulate \( p_{fib}(t) \) we truncate to \( t/T \in [-M, M] \). \( M \) should be chosen large enough such that \( p_{fib}(t) \approx 0 \) for \( t \not\in [-mT, mT] \), but not too large because we need to be able to represent the \( h_{fib}(t) \) oscillation at an instantaneous frequency of \( f(t) = \frac{\pi \gamma}{T_s^2} \). To meet the conditions of the Nyquist-Shannon sampling theorem for an arbitrary \( M \), we need \( 2N_s = 2 \frac{T_s}{f_{symbol}} = 2 \frac{f_s}{f_{symbol}} \geq \frac{\pi}{\gamma} M \). If we limit ourselves to \( \gamma \geq 1 \) and set \( N_s = T/T_s = 64 \), \( M = 10 \) is the best we can do without suffering from frequency aliasing. This \( M \) is fine for \( \gamma = 1 \) but is inadequate for higher \( \gamma \) where the sinc declines more slowly. An easy solution is to choose \( M(\gamma) = \left\lfloor \frac{N_s \gamma}{2\pi} \right\rfloor \).
2.2.2 Simulation of the BER

To evaluate our system we want to compare BERs before and after equalization. Calculating the BER analytically is only feasible for the most basic systems (e.g. the unipolar NRZ over the linear channel we described in Section 2.1.1) so we should resort to simulations.

The most straightforward way to simulate the BER is by doing a Monte Carlo experiment: continuously simulate the communication system at a given noise level until a certain amount of bit errors $K_e$ have occurred. The estimated BER is then given by \((\text{number of errors})/(\text{number of bits transmitted})\). To calculate the relative accuracy of the estimated BER we define:

$$
\epsilon^2 = \frac{\text{Var}(\text{BER}_{\text{est}})}{\text{BER}^2}
$$

(2.21)

It can be shown [19] that

$$
\epsilon^2 \approx \frac{1}{K_e}
$$

(2.22)

In order to get a close estimate of the BER, $K_e$ should at least be 100 and we should simulate $K_e/\text{BER}$ data bits. As we have to simulate a lot of BER curves and are also interested in low BER, this would take a very long time.

An alternative is evaluating the BER semi-analytically in three steps:

**Step 1** Simulate $K_{\text{data}}$ bits in the system without noise and determine the distance of the received signal to the decision threshold 0.5 as $e(k) = (2a(k) - 1)r(k)$ with $a(k)$ the original data symbol (0 or 1) and $r(k)$ the corresponding value at the input of the decision unit.

**Step 2** Determine the noise statistics at the input of the decision unit by simulating $K_{\text{noise}}$ symbol periods of noise in the system without data symbols and without offset in the equalizer. We know that $\mu_{\text{noise}}$ will be $\approx 0$ and calculate

$$
\sigma_0^2 = \frac{1}{K_{\text{noise}} - 1} \sum_{k=1}^{K_{\text{noise}}} |r(k) - \mu_{\text{noise}}|^2 = \frac{1}{K_{\text{noise}} - 1} \sum_{k=1}^{K_{\text{noise}}} |r(k)|^2
$$

(2.23)

**Step 3** Calculate the error probability for every bit analytically using the above two simulations as $P_{\text{error}}(k) = Q(e(k)/\sigma)$ and average over $k$ to obtain the BER.
The advantage of this method is that $K_{data}$ and $K_{noise}$ can be chosen independently of the BER value we expect to obtain, in our simulations we use $K_{data} = K_{noise} \geq 10^5$. A disadvantage is that not all systems can be evaluated with this method. When, for example, the decision of a symbol depends on the decisions of previous symbols (e.g. in Viterbi equalization), we cannot use this semi-analytical method.
Chapter 3

Discrete MMSE equalizer

This chapter gives an overview of the different configurations that we considered to equalize the sampled voltage \( r(k) \) at the TIA output. We always start from the same criterion: we minimize the mean square error (MSE) between the output of the equalizer and the data symbols \( a(k) \in \{0,1\} \). For each configuration we will investigate the behavior of the MSE and performance of the corresponding BER. In the BER plots we will use the BER curve for the system without dispersion and without equalization as a reference curve. To have an indication of the performance of the equalizer we will also plot the BER curve for the system with dispersion, but without equalization.

To analyze the performance of the MMSE equalizer we will mainly concentrate on four values of \( \gamma \). First we take \( \gamma = 1 \) for the case with very little dispersion. At \( \gamma = 4 \) the dispersion is not negligible anymore, but the eye in the eye diagram is still open. Finally we select \( \gamma = 6 \) and \( \gamma = 7 \) because for these values the eye is closed and a BER floor arises. In Figure 3.1 we plot the BER for these values of \( \gamma \).

3.1 Linear equalizer

First we try to equalize the non-linear system by means of a linear equalizer. In this configuration we neglect that our system in non-linear and try to equalize the system as if it is a linear one. The system that we consider is presented in Figure 3.2. We denote the output of the equalizer as \( u(k) \). To calculate \( \hat{a}(k) \) we use a simple symbol-by-symbol detector with \( u(k) \) as input. Since
3.1 Linear equalizer

Figure 3.1: BER curves for different $\gamma$.

our data bits are $a(k) \in \{0, 1\}$, we just decide if the output of the equalizer is larger than a decision threshold $d$.

Figure 3.2: Diagram of system with linear equalizer.

3.1.1 Algorithm

The algorithm that we use minimizes, as previously mentioned, the MSE:

$$MSE = E[|u(k) - a(k)|^2]$$

The problem reduces to finding equalizer coefficients $h_{FF}(m)$ and offset $c_{off}$ that minimize equation 3.1.

In the minimization we use the signal $r(k)$ rather than $s(k)$ because in $r(k)$ the noise contribution is also included. As a result the algorithm will take the noise into account, which leads to less noise enhancement at low values of $E_b/N_0$.

To perform the calculation of the equalizer coefficients we assume that the equalizer filter is finite with a length of $N_{FF}$ taps. $N_{FF,L}$ represents the number of non-causal taps and $N_{FF,R}$
the number of causal taps. The expression we have to minimize thus becomes:

\[ \text{MSE} = \mathbb{E} \left[ |u(k) - a(k) - c_{off}|^2 \right] \]  

(3.2)

with \( u(k) = \sum_{m=-N_{FF,L}}^{N_{FF,R}} r(k - m) h_{FF}(m) - c_{off} \)

To minimize expression 3.3 we equate the derivative of the MSE to the unknown quantities to zero. This means we have to calculate \( \frac{\partial \text{MSE}}{\partial h_{FF}(m)} \) and \( \frac{\partial \text{MSE}}{\partial c_{off}} \) \( \forall m \). This expression will result in a system of linear equations that can be expressed as:

\[ Rh_{FF} = r \Rightarrow h_{FF} = rR^{-1} \]  

(3.3)

Here we define the \( N_{FF} \times N_{FF} \) matrix \( R \) with \( R(i,j) = \mathbb{E}[r(k - m)r(k - m')] - \mathbb{E}[r(k - m)]\mathbb{E}[r(k - m')] \) with \( m = i - N_{FF,L} - 1 \) and \( m' = j - N_{FF,L} - 1 \) and vector \( r \) with \( r(i) = \mathbb{E}[a(k)r(k - m)] - \mathbb{E}[a(k)]\mathbb{E}[r(k - m)] \), again with \( m = i - N_{FF,L} - 1 \). In this expression the vector \( h_{FF} \) contains the filter coefficients. To obtain \( c_{off} \) we can plug in the values of \( h_{FF} \) in the next equation:

\[ c_{off} = \sum_{m=-N_{FF,L}}^{N_{FF,R}} h_{FF}(m) \mathbb{E}[r(k - m)] - \mathbb{E}[a(k)] \]

The expectation of the data symbols, \( \mathbb{E}[a(k)] \), is not difficult to calculate. As we send 0 or 1 and we assume that the data symbols are a result of an independent random process, we get \( \mathbb{E}[a(k)] = 1/2 \). The other expectations are rather difficult to calculate analytically and hence we simulate them.

Our algorithm calculates the filter coefficients that result in a MMSE. This is not the same as the filter coefficients that result in minimum BER, but it can be expected that an improvement in MSE will usually lead to a lower BER. [20]. At the end of this chapter, however, we present a simple counterexample.

To calculate the MSE we use the expression:

\[ \text{MSE} = \frac{1}{4} - r^T R^{-1} r \]  

(3.4)

that we derive in Appendix C.1.
3.1 Linear equalizer

3.1.2 Simulation results

First remark that to equalize the system, we have to decide when we sample the output signal \( r(t) \) to obtain \( r(k) \). The signal \( r(k) \) is the sampled version of \( r(t) \) at a sample rate of \( 1/T \). This means that \( r(k) = r(kT + \tau) \) with \( 0 \leq \tau < T \). So to sample the signal \( r(kT + \tau) \) we must determine the best \( \tau \). The obvious choice is to sample at the time instant corresponding to the largest eye opening, this is also the best moment if you do not perform any equalization. With a linear equalizer however this is not necessarily true. We equalize to get the minimum MSE so we also use this criterion to determine \( \tau \). We apply the algorithm on the different \( N_s \) possible sample moments in our simulation and select the one resulting in the lowest MSE.
Figure 3.3: Eye diagram and MSE for $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$. 
First we take a look at the minimal MSE for $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$ for an equalizer with 11 taps. In Figure 3.3 we present the eye diagram of $r(t)$ and the associated minimal MSE as function of the sample moment. With the dashed vertical line we indicate the sample moment corresponding to the largest eye opening. The dotted vertical line indicates the moment of minimum MSE. Note that these lines do not coincide. On the eye itself we mark with the thick dashed line the smallest values of the waveforms corresponding to the symbol 1. Similarly we mark with the thick dash-dotted line the largest values of the waveforms corresponding to the symbol 0. The eye is by definition the region between those two curves.

Further it is clear that the MSE increases for increasing values of $\gamma$. This is as expected, since more dispersion is added for higher values of $\gamma$.

Next note that for all values of $\gamma$ the MSE after equalization is lower than the MSE without equalization. This is of course logical, since our algorithm tries to minimize the MSE.

Finally we observe that in the case of $\gamma = 1$ the eye opening is rather broad, which results from the fact that there is only little dispersion. Due to this low dispersion the MSE shows a rather broad minimum over the whole eye opening. This means that the exact sample moment does not really matter to obtain the minimum MSE. For $\gamma = 4$ the eye width is already smaller and that is also visible in the figure of the MSE where the minimum is less broad. At $\gamma = 7$ the eye is closed, so we dropped the dashed vertical line. Nevertheless it is possible for the algorithm to find coefficients that lower the MSE. Here the minimum is sharp, meaning that the sample moment has to be more accurate than for $\gamma = 1$.

In the previous figures we applied a filter with 11 taps ($N_{FF,L} = N_{FF,R} = 5$). Of course we expect that the performance of the equalizer filter and hence the MSE will depend on the number of taps. To check its influence we plot in Figure 3.4 the MSE as function of the sample moment for different numbers of filter taps for both $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$. It is logical that the MSE decreases if the number of taps increases, since the equalizer becomes longer and hence has more degrees of freedom. However, we have to remark that the decrease in MSE becomes insignificant for longer filter lengths. Only the increase from 1 tap to 3 taps corresponds to a significant decrease in MSE. This is of course because a 1 tap equalizer only scales and offsets the signal.
3.1 Linear equalizer

Figure 3.4: MSE as function of the sample moment for different number of filter taps for $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$. 
Next we take a look in Figure 3.4 at the BER performance of the system with $\gamma = 1$. As can be seen on the top figure, the influence of the equalizer is rather limited which is as expected since not much dispersion is added. The effect of changing the length of the filter or the location of the taps (causal or anti-causal) is almost not visible. As increasing the filter length has little effect, we conclude that compensating about one fourth of the degradation due to dispersion is the best we can do in this case with a linear equalizer.

![Figure 3.5: BER curves for $\gamma = 1$.](image)

In Figure 3.6 we plot the BER for different lengths of the equalizer filter with $\gamma = 4$. We included equalizers with only causal or anti-causal taps to show that they perform differently, which is to be expected because the TIA does not have a symmetric impulse response. From Figure 3.6 we can conclude that the length of the equalizer filter does not really matter at $\gamma = 4$. At a BER level of $10 \times 10^{-5}$ we get a gain of a bit more than 2 dB in terms of $E_b/N_0$, but this is still not half of the degradation added by the dispersion.

Finally we take a look at the BER of the system with high dispersion where the eye diagram before equalization is closed. This is the case for both $\gamma = 6$ and $\gamma = 7$. The BER curves can be found in Figure 3.7.

For $\gamma = 6$ we see that without equalization we have a BER floor of about $6 \times 10^{-3}$. This floor is relatively low because for $\gamma = 6$ the eye is closed but only a few of the possible data sequence will result in a sample that lies in the wrong decision region. On the figure we notice that this BER floor already disappears with a short equalizer. Note that the BER only starts to go down quickly for quite high values of $E_b/N_0$, we suppose that this is because compensating a
3.1 Linear equalizer

Figure 3.6: BER curves for $\gamma = 4$.

High dispersion results in larger noise enhancement and hence the sharp decrease of the BER will be at higher $E_b/N_0$. Note that increasing the filter length here provides an extra gain of approximately 2 dB at high $E_b/N_0$. This is not unexpected as the higher dispersion will result in more broadening of the pulse and hence more ISI will be present at the entrance of the PD and more symbols will have an influence on the current sample.

For $\gamma = 7$ we can conclude from the figure that, even for very long filters, the linear equalizer lowers the BER floor but does not solve it. This shows that with a linear equalizer filter we only can eliminate a BER floor if the dispersion is not too high.
3.1 Linear equalizer

We suspect that the reason we cannot remove the BER floor is the non-linearity in the PD which results in quadratic terms. We verify this by simulating the system where we replace the PD $| \cdot |^2$ by $\text{Re}( \cdot )$. Figure 3.8 shows that the linear system at $\gamma = 7$ can be equalized without any problem even though the BER floor without equalization is approximately the same as in
the non-linear system. In Figure 3.9 we show the influence of the filter length on the linear system performance, increasing from $N_{FF,L} = N_{FF,R} = 10$ to $N_{FF,L} = N_{FF,R} = 20$ has almost no effect on the BER.

Figure 3.8: Influence non-linearity on equalizer performance.

Figure 3.9: Influence filter length on linear system.
3.2 Fractionally spaced linear equalizer

3.2.1 Concept

In Chapter 2 we derived the following expression for the data component of $r(t)$:

$$s(t) = \sum_{m} a(m) \sum_{k=-M}^{M} a(m+k)h_{tot,k}(t-mT)$$  \hspace{1cm} (3.5)

We interpret this as the superposition of $2M+1$ transmit pulses $h_{tot,k}(t)$ convolved with $2M+1$ data sequences $a(m)a(m+k)$, $k \in [-M,M]$. Of these $2M+1$ components of $s(t)$, we are only interested in the one corresponding to $k = 0$. The idea of taking multiple samples per second is that we have extra degrees of freedom to not only equalize the linear terms, but to also suppress the extra terms due to the non-linearity. Schematically the system that we now consider is presented in Figure 3.10.

Figure 3.10: Diagram of system with fractional linear equalizer.

To show that the linear fractional equalizer results in more freedom we will first investigate the situation of a linear system. We show that a zero forcing (ZF) equalizer of finite length can be found if enough samples per symbol are provided. The fractionally spaced equalizer can be interpreted as the linear combination of $M_s$ different equalizers, each working at a symbol rate $1/T$. Because every one of the $M_s$ samples is taken at a different sample moment, they can be seen as being sent through $M_s$ different channel filters. Figure 3.11 shows this structure schematically. $H_i(z)$ is the Z-transform of the $i^{th}$ channel response, whereas $H_{eq,j}(z)$ represents the Z-transform of the $j^{th}$ equalizer filter. To obtain no ISI in the signal $u(k)$, the following equation must hold:

$$\sum_{m=1}^{M_s} H_{eq,m}(z)H_m(z) = 1$$  \hspace{1cm} (3.6)

If we limit the length of all equalizers to $N_{eq}$ coefficients and assume that the channel filters have $N$ coefficients, the product $H_{eq,m}(z)H_m(z)$ will have a length of $N_{eq} + N - 1$. As a result we will have $N_{eq} + N - 1$ linear equations with $M_sN_{eq}$ unknowns and to find a solution inequality 3.7 must be true.

$$N_{eq} + N - 1 \leq M_sN_{eq}$$  \hspace{1cm} (3.7)
Now we turn our attention to the non-linear case. Due to the square in the PD and the ISI at its entrance, we will get quadratic terms of the data symbol $a(k)$ with a shifted version $a(k - 1)$ as can be seen in equation 3.5. Let $a(k)a(k - j) = a_j(k)$ with $j = 0, 1..J - 1$, then we can interpret the signal as a sum of $J$ data streams $a_j(k)$, each filtered by a filter $h_{tot,k}(t)$. In Figure 3.12 we present the corresponding structure of the fractionally spaced equalizer. Note that this structure is obtained by the same reasoning as in the linear case. The difference is that due to the non-linearity we now have multiple data streams $a_j(k)$. Since we are only interested in the data stream of $a_0$, the ZF equalizer condition becomes:

$$\sum_{m=1}^{M_s} H_{eq,m}(z)H_{m,j}(z) = \delta_j$$ (3.8)

In equation 3.8, $H_n,j(z)$ is the Z-transform of the sampled $h_{tot,k}(t)$ and $H_{eq,m}(z)$ the Z-transform of the $m^{th}$ equalizer. If we search for solutions with equalizers of length $N_{eq}$ and assume $H_{n,j}(z)$ has $N$ coefficients, we get $J(N + N_{eq} - 1)$ linear equations, with $N_sN_{eq}$ unknowns. A necessary condition for the existence of a solution is:

$$J(N_{eq} + N - 1) \leq M_sN_{eq}$$ (3.9)

This means that for each $J$ there exists a value $M_s$ that fulfils equation 3.9. Note that $J$ is theoretically infinite for our pulse $p_{fib}$. In practice however, $p_{fib}(t)$ becomes negligible for large $|t|$. To check if the above reasoning is correct, we will apply the fractionally spaced equalizer to an artificial pulse in the simple system described in Section 3.2.3.

Note that this reasoning is done for the ZF equalizer, whereas in the simulations we use the MMSE equalizer. If the BER floor can be removed by a ZF equalizer however, this is also possible for the MMSE equalizer as they converge to each other for $E_b/N_0$. 

---

**Figure 3.11:** Linear system with fractional equalizer.
3.2 Fractionally spaced linear equalizer

3.2.2 Algorithm

We again try to minimize the MSE. The calculation of the filter coefficients is similar as in the case of linear equalization, we refer to Appendix C.2 for the complete calculation. We again obtain a system of linear equations, now $R(i, j) = E[r(kM_s + \tau - m')r(kM_s + \tau - m)] - E[r(kM_s + \tau - m')]E[r(kM_s + k_0 - m)]$ with $m = i - N_{FF,L} - 1$ and $m' = j - N_{FF,R} - 1$ and $r(i) = E[a(k)r(kM_s + \tau - m)] - E[a(k)]E[r(kM_s + \tau - m)]$.  

Figure 3.12: Non-linear system with fractional equalizer.
3.2 Fractionally spaced linear equalizer

3.2.3 Simple example

To test our concept and to gain more insight, we consider a channel with an artificial pulse $p_{\text{alt}}(t)$. Since in this section we only test if our concept is correct, we limit the model to the most necessary components of our system: we drop the TIA and replace it by a perfect analog filter with a cut-off frequency equal to $f_s/2$, the sample frequency of our simulation. This means that in the simulation we can replace the TIA by a Dirac impulse. As the TIA is a linear filter after the PD non-linearity, this adaptation will not influence the main operation of the algorithm. Of course this change will increase the noise power at $r(t)$ because the TIA filtered the noise at approximately $1/T$, whereas the new filter filters at $f_s/2$. This is why we simulate at higher signal levels. The system is presented in Figure 3.13.

![Figure 3.13: Diagram of simple system.](image)

To keep it simple, we will limit the alternative pulse $p_{\text{alt}}(t)$ in time. The pulse that we consider is given in equation 3.10 where we use the parameter $\gamma$ to indicate the length of the pulse. The parameter $\beta$ controls the frequency of the oscillation and is held at 1. The decay in time is handled by parameter $\mu$. The larger $\mu$, the faster the pulse will extinguish. To make sure the pulse is finite in time we use the window $w(t)$ presented in equation 3.11. The pulse is, just as our $p_{\text{fib}}(t)$, symmetrical around $t = 0$. To avoid normalization issues we choose $T = 1$. This will not influence the system performance.

$$p_{\text{alt}}(t) = \begin{cases} \cos \left( \frac{\pi \beta t}{T} \right) \exp(-\mu|t|)w\left(\frac{t}{T}\right) & -\gamma T \leq t < \gamma T \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (3.10)

with $w(t) = \begin{cases} 1 & -\gamma + 1 \leq t \leq \gamma - 1 \\ \sin \left( \left( t + \gamma \right) \frac{\pi}{2} \right) & -\gamma \leq t < -\gamma + 1 \\ \sin \left( \left( -t + \gamma \right) \frac{\pi}{2} \right) & \gamma - 1 \leq t < \gamma \\ 0 & \text{otherwise} \end{cases}$  \hspace{1cm} (3.11)

In Figure 3.14 we plot the pulse $p_{\text{alt}}(t)$ together with the window $w(t)$. Note that, if we sample the signal at multiples of the symbol period $T$, there is influence from $\gamma$-1 prior symbols and
γ-1 symbols after the symbol we consider (black dots). The pulse that we consider here is not related to any physical structure whatsoever, to perform this analysis we also could truncate the pulse $p_{fib}(t)$. The advantage of $p_{alt}(t)$ is however its closed-form analytical expression.

![Figure 3.14: $p_{alt}(t)$ for $\gamma = 2$, $\mu = \ln(1/2)$ and $\beta = 1.$](image)

To understand the behavior of the algorithm, we study histograms of the error vector at the input of the decision unit. We will consider both the system with and without equalization yielding, respectively $u(k)$ and $r(k)$. The error is defined as

$$e(k) = (a(k) - u(k))(2a(k) - 1) = \begin{cases} 
    u(k) & \text{if } a(k) = 0 \\
    1 - u(k) & \text{if } a(k) = 1
\end{cases}$$

(3.12)

and analogous for the system without equalization (replace $u(k)$ by $r(k)$). This way positive $e(k)$ correspond to errors towards the decision boundary and negative $e(k)$ to errors away from the decision boundary, this is shown schematically in Figure 3.15. As a result all $e(k) > 0.5$ lead to an erroneous decision and all $e(k) < 0.5$ to a correct decision. We present the histograms for the samples corresponding to $a(k) = 0$ and $a(k) = 1$ separately wherever this is of interest. It is interesting to study the histograms of the system without noise because the occurrence of $e(k) > 0.5$ corresponds with a BER floor.

We will investigate the system with $\gamma = 2$, as a result $r(k)$ we will contain influence of the symbols $a(k - 1)$, $a(k)$ and $a(k + 1)$ if we sample at sample moment $kT$. The gain of symbol $a(k)$ is equal to one, since $p_{alt}(0) = 1$ and due to the symmetry the gain for symbols $a(k - 1)$ and $a(k + 1)$
and \( a(k + 1) \) is the same and equal to \( \exp(-\mu) \). Let \( A \) be that gain, then we get:

\[
    r(k) = |Aa(k-1) + a(k) + Aa(k+1)|^2
    = A^2a(k-1)^2 + a(k)^2 + A^2a(k+1) + 2Aa(k-1)a(k) + 2Aa(k)a(k+1) + 2A^2a(k-1)a(k+1)
\]

This means that due to the spreading of the pulse and the square, there are \( 2^{2\gamma-1} \) possible levels in \( r(k) \), they are presented in Table 3.3. Due to the symmetry of the pulse these levels correspond not to eight, but six different values.

| Table 3.1: Possible values of \( r(k) \) for \( \gamma = 2 \). |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( a(k-1) \)    | \( a(k) \)      | \( a(k+1) \)    | \( r(k) \)      | \( e(k) \)      |
| 0               | 0               | 0               | 0               | 0               |
| 1               | 0               | 0               | \( A^2 \)       | \( A^2 \)       |
| 0               | 0               | 1               | \( A^2 \)       | \( A^2 \)       |
| 1               | 0               | 1               | \( 4A^2 \)      | \( 4A^2 \)      |
| 0               | 1               | 0               | 1               | 0               |
| 1               | 1               | 0               | \( (A+1)^2 \)   | \( 1 - (A+1)^2 \) |
| 0               | 1               | 1               | \( (A+1)^2 \)   | \( 1 - (A+1)^2 \) |
| 1               | 1               | 1               | \( (2A+1)^2 \)  | \( 1 - (2A+1)^2 \) |

Both \( a(k) = 1 \) and \( a(k) = 0 \) lead to three possible levels of \( r(k) \). It is clear that if one of these levels is beyond the decision threshold \( d \), we will have a BER floor. If we for example let \( (2A + 1)^2 < 1/2 \) then \( A \) should be between \( \frac{-1 - \sqrt{2}}{2} \) and \( \frac{-1 + \sqrt{2}}{2} \). Every \( A \) between those limits and \( 1/2 \) as decision threshold will result in a BER floor.

First we will investigate the case of \( A = -1/5 \). The possible values for \( r(k) \) are listed in Table 3.2. Since only one of the eight possibilities lies in the wrong decision region, we expect a
Table 3.2: Possible values of $r(k)$ for $\gamma = 2$ and $A = -1/5$.

<table>
<thead>
<tr>
<th>$a(k-1)$</th>
<th>$a(k)$</th>
<th>$a(k+1)$</th>
<th>$r(k)$</th>
<th>$e(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1/25</td>
<td>1/25</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1/25</td>
<td>1/25</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4/25</td>
<td>4/25</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>16/25</td>
<td>9/25</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>16/25</td>
<td>9/25</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9/16</td>
<td>16/25</td>
</tr>
</tbody>
</table>

BER of 1/8. This is confirmed by the line with $d = 1/2$ in Figure 3.16 where we plot the BER of the system with $A$ equal to -1/5. This error floor however is avoidable, since it can easily be solved by changing the decision threshold $d$. In Table 3.2 we see that there is no overlap between the values originating from symbol 0 and 1: the highest value corresponding with the symbol 0 is 4/25 and the lowest value corresponding with the symbol 1 is 9/25. Hence the most obvious choice for $d$ is 13/50. In Figure 3.16 we see that this choice for the decision threshold indeed solves the error floor.

![Figure 3.16: BER for $\gamma=2$ and $\mu = \ln(5)$.](image)

Now we consider a second case where $A = -1/2$ and hence $\mu = \ln(2)$. The different values for $r(k)$ are listed in Table 3.3.
Table 3.3: Possible values of $r(k)$ for $\gamma = 2$ and $A = -1/2$.

<table>
<thead>
<tr>
<th>$a(k-1)$</th>
<th>$a(k)$</th>
<th>$a(k+1)$</th>
<th>$r(k)$</th>
<th>$e(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1/4</td>
<td>1/4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1/4</td>
<td>1/4</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1/4</td>
<td>3/4</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1/4</td>
<td>3/4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Here four out of the eight possible bit sequences will result in a value in the wrong decision region, which results in an error floor of 1/2. Further, due to the choice of $A$, we only get three different values for $r(k)$: 0, 1/4 and 1. The real problem is however that for each value the probabilities $p(0|r(k)) = p(1|r(k)) = 1/2$. Hence the posterior probability is the same as the prior probability because observation of $r(k)$ will not give us any extra information.

First we try to equalize this system with one tap, to see what happens. The advantage of using one tap is that the equalizer filter is in fact just a scaling factor $h$ and an offset $c$. In this simple case we can calculate these factors $h$ and $c$ for a given $A$:

$$h = \frac{2A + 1}{9A^4 + 8A^3 + 8A^2 + 4A + 1}$$ (3.13)

$$c = \frac{-1}{2} \frac{A^2(9A^2 + 2A + 1)}{9A^4 + 8A^3 + 8A^2 + 4A + 1}$$ (3.14)

If we plug in $A$ equal to -1/2 in these equations we get $h = 0$ and $c = -1/2$, which will result in an equalized signal $u(k) = 1/2$ and an error floor of 1/2. It is logical that in this situation we cannot equalize the signal, since we cannot tell anything from the observation $r(k)$. Increasing the length of the filter has also no effect at all. The BER curves can be found in Figure 3.17.

Finally we investigate the more interesting case of $A = -3/5$ and $\mu = \ln(5/3)$. The different values for $r(k)$ are given in Table 3.4. If we put the decision threshold $d$ at 1/2, we expect a BER of 0.5. Note that we indeed have an unavoidable error floor since we can not solve it by changing the decision threshold. This is because the highest value corresponding to the symbol 0 is larger that all the values corresponding to the symbol 1. On the histogram of the error
3.2 Fractionally spaced linear equalizer

Figure 3.17: BER for $\gamma = 2$ and $\mu = \ln(2)$.

presented in Figure 3.18, this can be clearly seen.

Table 3.4: Possible values of $r(k)$ for $\gamma = 2$ and $A = -3/5$.

<table>
<thead>
<tr>
<th>$a(k-1)$</th>
<th>$a(k)$</th>
<th>$a(k+1)$</th>
<th>$r(k)$</th>
<th>$e(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$9/25$</td>
<td>$9/25$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>$9/25$</td>
<td>$9/25$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$36/25$</td>
<td>$36/25$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$4/25$</td>
<td>$21/25$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>$4/25$</td>
<td>$21/25$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$1/25$</td>
<td>$24/25$</td>
</tr>
</tbody>
</table>

First we take a look at what happens if we try to equalize $r(k)$ with one tap and one sample per symbol. If we plug in $A = -3/5$ in equations 3.13 and 3.14 we get $h = -0.2178$ and $c = -0.5958$. Due to the minus sign in $h$, the error will change sign and now we will detect three out of the four possibilities of symbol 0 and only one possibility of symbol 1 erroneously. This can be seen in the histogram in Figure 3.19.

We found that a ZF equalizer for a non-linear channel can be found if the inequality $J(N_{eq} + N - 1) \leq M_s N_c g$ is fulfilled. As $p_{alt}$ has a length of $4T$, every symbol interval of $r(t)$ contains contributions of four symbols, i.e. $N = 4$. The PD leads to contributions of the type $a_l(k) =$
3.2 Fractionally spaced linear equalizer

\[ a(k)a(k-l) \] with \( l \in [0, 1, 2, 3] \), i.e. \( L = 4 \). If our concept is correct we should be able to equalize this system with 8 samples per symbol interval and an equalizer filter that uses the previous and next symbol interval (\( N_{eq} = 3 \)). To verify this, we plot in Figures 3.20 - 3.22 the histograms of the error vector for increasing values of \( M_s \).

In the histograms we can see that for \( M_s = 2 \) and \( M_s = 4 \) the error floor is not gone. Nevertheless we see that the error floor will be lower since, especially in the case of \( M_s = 4 \), there are already more errors that are lower than 0.5. For \( M_s = 8 \) the histograms shows us that in this case the

Figure 3.18: Histogram of the unequalized system for \( \gamma = 2 \) and \( \mu = \ln(5/3) \).

Figure 3.19: Histogram of the equalized system with one tap for \( \gamma = 2 \) and \( \mu = \ln(5/3) \).
BER floor disappears and the equalization works as predicted. These observation are confirmed by the BER in Figure 3.23. The BER floor drops to approximately 0.4 for $M_s = 2$ and 0.1 for $M_s = 4$.  

Figure 3.20: Histogram of the equalized system with $M_s = 2$ for $\gamma = 2$ and $\mu = \ln(5/3)$.

Figure 3.21: Histogram of the equalized system with $M_s = 4$ for $\gamma = 2$ and $\mu = \ln(5/3)$.  

3.2 Fractionally spaced linear equalizer
3.2 Fractionally spaced linear equalizer

Figure 3.22: Histogram of the equalized system with $M_s = 8$ for $\gamma = 2$ and $\mu = \ln(5/3)$.

Figure 3.23: BER for $\gamma = 2$, $\mu = \ln(5/3)$ and different $M_s$.

3.2.4 Simulation results

Now we return to our system with the complex channel pulse. For the linear equalizer we had $N_s$ possible sample moments in the simulation, since we simulate the continuous time signals at a sample rate of $N_s/T$. This is analogous in the case of the fractionally spaced equalizer. Now we first sample the continuous signal at a sample rate of $M_s/T$ with $M_s$ the number of samples per symbol, equalize this signal with the fractionally spaced equalizer and then downsample the result with a factor $M_s$. 
The equalizer now works $M_s$ times faster than in the case of one sample per symbol. This means that a filter with the same length in time now has $M_s$ more taps. The number of taps in the figures and tables is always the absolute number, not the number of symbol periods. To have a comparison based on the absolute equalizer length, one should divide the number of taps by $M_s$.

Again we first investigate the minimum MSE that we can achieve. To do so we list in Table 3.5–3.7 the minimum MSE for $\gamma = 1$, $\gamma = 4$ and $\gamma = 7$ respectively. We expect that the longer we choose the filters, the lower the MSE will become because the equalization freedom increases. The same reasoning of course applies to $M_s$. It is clear in the table that the MSE decreases for increasing $M_s$ and increasing lengths as expected.

Table 3.5: MSE in case for $\gamma = 1$ and multiple samples per symbol.

<table>
<thead>
<tr>
<th>$M_s$ = 2</th>
<th>$M_s$ = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>filter length</td>
<td>MSE</td>
</tr>
<tr>
<td>1</td>
<td>5.6E-4</td>
</tr>
<tr>
<td>5</td>
<td>6.1E-5</td>
</tr>
<tr>
<td>21</td>
<td>7.9E-6</td>
</tr>
</tbody>
</table>

Table 3.6: MSE in case for $\gamma = 4$ and multiple samples per symbol.

<table>
<thead>
<tr>
<th>$M_s$ = 2</th>
<th>$M_s$ = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>filter length</td>
<td>MSE</td>
</tr>
<tr>
<td>1</td>
<td>1.3E-2</td>
</tr>
<tr>
<td>5</td>
<td>9.7E-4</td>
</tr>
<tr>
<td>21</td>
<td>3.0E-4</td>
</tr>
</tbody>
</table>

Table 3.7: MSE in case for $\gamma = 7$ and multiple samples per symbol.

<table>
<thead>
<tr>
<th>$M_s$ = 2</th>
<th>$M_s$ = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>filter length</td>
<td>MSE</td>
</tr>
<tr>
<td>1</td>
<td>1.0E-1</td>
</tr>
<tr>
<td>5</td>
<td>6.2E-2</td>
</tr>
<tr>
<td>21</td>
<td>3.0E-2</td>
</tr>
</tbody>
</table>
To see the effect of the equalization at a higher rate, we present in Figure 3.24 the histograms of the error made at $u(k)$ for $\gamma = 1$. The same is done in Figure 3.25 for $\gamma = 4$. In both cases the equalizer filter is 10 symbol periods long and no noise is added. For these values of $\gamma$ the eye is still open. This can be seen on the figures as there are no errors larger than $1/2$. In the histograms we observe that the equalizer reduces the error with increasing $M_s$ and hence the histograms are more centred around zero. If the noise enhancement due to the equalizer filter is not too high, the noise margin will increase and the corresponding BER decreases.

![Figure 3.24: Histogram of the error for $M_s = 1$(left), 2(center) and 4(right) for $\gamma = 1$.](image1)

![Figure 3.25: Histograms of the errors for $M_s = 1$(left), 2(center) and 4(right) for $\gamma = 4$.](image2)

In Figure 3.26 we plot the histograms for $\gamma = 7$. The difference with the previous figures is that for this $\gamma$ the eye is closed and a BER floor arises. This can be seen on the histograms by the fact that there are errors larger than $1/2$ (indicated by the vertical line). The influence of increasing $M_s$ is visible here too. Just as in the previous case, we see that the histogram becomes centered around 0. The only problem is that with these values of $M_s$, the BER floor becomes smaller (less errors larger than one half) but does not disappear. We expect however, as our concept and the simple example predict, that the BER floor will disappear for increasing $M_s$.

To verify if this reasoning is correct, we make in Figure 3.27 the histogram for $\gamma = 7$, $M_s = 32$ and a filter that is 20 symbol periods long. This means that we have a filter of 641 taps, which results in a large computational complexity and is hence not very practical. From the histogram however we can conclude that the BER floor disappears in this simulation with $1 \times 10^6$ symbols and no noise.

Figure 3.28 shows a plot of the BER for different values of $M_s$ with $\gamma = 1$. In this case of low dispersion it is possible to achieve the same and even a slightly better BER than in the case
of no dispersion. This means that the equalizer completely removes the effect of the dispersion and even compensates for other smaller imperfections such as the ISI from the TIA. Remark that in the right figure, where we zoomed in on the BER for high $E_b/N_0$, we see that the BER for $M_s = 16$ is worse than the BER for $M_s = 4$. One explanation for this is that the simulation we performed is not accurate enough. Another explanation might be that the MSE is lower, but the corresponding BER is not.

Figure 3.28: BER as function of $E_b/N_0$ with $\gamma = 1$ and $M_s = 1, 2, 4, 8$ and 16.
In Figure 3.29 we see the plot of the BER for $\gamma = 4$ and $M_s = 1, 2, 4, 8$ and 16. Here we see that the increase from $M_s = 1$ to 2 corresponds to a small gain, approximately 1 dB at high $Eb/N_0$, but that the gain for going to even larger values of $M_s$ is almost non-existent. This is clear on the right figure where we zoom in on the higher $Eb/N_0$.

Figure 3.29: BER as function of $Eb/N_0$ with $\gamma = 4$ and $M_s = 1, 2, 4, 8$ and 16.

The BER plot for $\gamma = 7$ is given in Figure 3.30. As predicted by the study of the MSE, the BER is limited by an error floor that is decreasing for increasing $M_s$. For the low values of $M_s$ we simulate however we cannot eliminate this error floor. From the histogram in Figure 3.27 we expect that the BER floor would disappear for $M_s = 32$ and a filter length of 20 symbol periods. In Figure 3.31 we present the corresponding BER curve. We immediately see that the error floor is not gone, but just lowered to approximately $1 \times 10^{-5}$. This low value also explains why this error floor was not visible on the histogram. This figure shows us that the BER floor keeps decreasing for increasing $M_s$. We expect that at one point this floor will be completely gone.

Figure 3.30: BER as function of $Eb/N_0$ with $\gamma = 7$ and $M_s = 1, 2, 4$
3.3 Decision feedback equalizer

3.3.1 Concept

Next we use the decision of previous symbols to equalize the current sample by adding a filtered version of the previous symbol estimates. This principle is called decision feedback. Now we expect that the linear forward filter $h_{FF}(m)$ will reduce the precursor ISI and the feedback filter $h_{FB}(m)$ will deal with the postcursor ISI.

Compared to the system with only a linear forward filter, we expect that $\sum m |h_{FF}(m)|^2$ in this system will be smaller since the the feedback filter will eliminate the postcursor ISI. The result is that, since the noise only goes through the forward filter, there will be less noise enhancement compared to the system with only a forward filter.

A possible disadvantage of decision feedback is that an erroneous estimate of the previous symbol will influence the decision of the current symbol. Hence the probability that the current estimate will be erroneous increases, which could result in error propagation.

The diagram of the system with decision feedback can be found in Figure 3.32. Remark that, in contrast to the forward filter, the feedback filter always operates at the symbol rate, independent of $M_s$. 

Figure 3.31: BER as function of $E_b/N_0$ with $\gamma = 7$ and $M_s = 4$ and 32
3.3 Decision feedback equalizer

3.3.2 Algorithm

Since we added an extra filter, the expression for the MSE changes:

\[
MSE = E[|z(k) - a(k)|^2] = E\left[\sum_m h_{FF}(m)r(kM_s + \tau - m) - \sum_{l>0} h_{FB}(l)\tilde{a}(k-l) - c_{off} - a(k)|^2\right]
\]

Even though the equalizer now no longer is linear, the calculation of the equalizer coefficients is again similar. We again get a matrix \( R \) and \( r \). For the exact calculations we refer to Appendix C.3.

3.3.3 Simulation results

Decision feedback with \( a(k) \)

In this section we first consider the system where we use the correct, original data symbols \( a(k) \) for both the calculation of the filter coefficient and the equalization in the feedback path. Doing so, we do not have feedback in the system, since we do not use the detected symbol. The advantage is that we can still simulate the BER with the semi-analytical method.

In Table 3.8 we present the MSE for the different configurations of the decision feedback for \( \gamma = 1 \ , \gamma = 4 \) and \( \gamma = 7 \). The \( M_s \) and number of taps in the feedback path are also varied. The forward filter is 10 symbol periods long for all cases. From the tables we conclude that the MSE decreases for increasing feedback filter length. However if we compare this decrease with the decrease in MSE that can be achieved by increasing \( M_s \), we see that increasing the number of samples per symbol has more effect.
Table 3.8: MSE different configuration with decision feedback.

<table>
<thead>
<tr>
<th>$M_s$ taps $h_{FB}$</th>
<th>MSE for $\gamma = 1$</th>
<th>MSE for $\gamma = 4$</th>
<th>MSE for $\gamma = 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6.4E-5</td>
<td>1.7e-3</td>
<td>5.4e-2</td>
</tr>
<tr>
<td>1</td>
<td>3.9E-5</td>
<td>1.6e-3</td>
<td>2.8e-2</td>
</tr>
<tr>
<td>5</td>
<td>3.8E-5</td>
<td>1.3e-3</td>
<td>1.9e-2</td>
</tr>
<tr>
<td>0</td>
<td>7.5E-6</td>
<td>3.0e-4</td>
<td>3.1e-2</td>
</tr>
<tr>
<td>2</td>
<td>7.2E-6</td>
<td>2.8e-4</td>
<td>2.0e-2</td>
</tr>
<tr>
<td>5</td>
<td>7.2E-6</td>
<td>1.4e-4</td>
<td>1.4e-2</td>
</tr>
<tr>
<td>0</td>
<td>3.7E-6</td>
<td>8.7e-6</td>
<td>1.7e-2</td>
</tr>
<tr>
<td>4</td>
<td>3.5E-6</td>
<td>8.0e-6</td>
<td>1.1e-2</td>
</tr>
<tr>
<td>5</td>
<td>3.3E-6</td>
<td>5.6e-6</td>
<td>9.2e-3</td>
</tr>
</tbody>
</table>

The histograms corresponding to Table 3.8 with $M_s = 1$ can be found in Figures 3.33, 3.34 and 3.35. In these figures the forward filter is always ten symbol periods long. In the left figures there is no feedback, in the middle figures the feedback filter has one tap and in the right figures the feedback filter is 5 taps long.

For $\gamma = 1$ and $\gamma = 4$ the influence of the decision feedback is barely noticeable on the histograms. The error gets a little bit more centered around zero, but not spectacularly.

For $\gamma = 7$ the histogram shows a more interesting behavior. Without equalization the eye is closed for this dispersion and a BER floor arises. If we apply a feedback filter however, we notice in the histogram that the probability of having an error larger than 0.5 is lower than $10^{-4}$. This could mean that the BER floor becomes really low or that at high $E_b/N_0$ the BER floor disappears completely. Note that, to make these histograms we simulated $4 \times 10^{-5}$ data symbols. This should be long enough to make conclusions about probabilities larger than $1 \times 10^{-4}$.

![Figure 3.33: Histograms of the error for $h_{FB}$ with 0,1 and 5 taps for $\gamma = 1$.](image)

For $\gamma = 1$ the plot of the BER is shown in Figure 3.36. In this plot we include the BER for $M_s = 1$ and $M_s = 4$, once with and once without feedback. We clearly see no influence of
3.3 Decision feedback equalizer

In practice this means that decision feedback is not interesting for low dispersion since it will only add complexity and does not amount to any real improvement in performance.

Figure 3.34: Histograms of the error for $h_{FB}$ with 0, 1 and 5 taps for $\gamma = 4$.

Figure 3.35: Histograms of the error for $h_{FB}$ with 0, 1 and 5 taps for $\gamma = 7$.

The decrease is the same for $M_s = 1$ and $M_s = 4$ and corresponds to a gain of approximately 0.5 dB at high $E_b/N_0$.

Figure 3.36: BER curve for $\gamma = 1$.

In Figure 3.37 we plot for $\gamma = 4$ the BER of the system with decision feedback and $M_s = 1$ or $M_s = 4$. Contrary to the results for $\gamma = 1$, the decision feedback now results in an improved BER.

Figure 3.38 contains the BER curves for $\gamma = 7$ with decision feedback. As predicted by the histogram, the BER floor disappears for $M_s = 1$, a forward filter of ten symbols periods long
3.3 Decision feedback equalizer

Figure 3.37: BER curve for $\gamma = 4$ for $M_s = 1$ (top figure) and for $M_s = 4$ (bottom figure)

and a feedback filter with only one tap. In the figure we also plot the BER for $M_s = 4$ and no
decision feedback. As mentioned before, we also expect the BER to disappear for increasing
$M_s$ but this corresponds to a much higher complexity than adding one tap in the feedback path.

To show that decision feedback will not eliminate the BER floor for all $\gamma$, we provide in
Figures 3.39 and 3.40 the BER curves for $\gamma = 8$ and $\gamma = 10$ respectively.
For $\gamma = 8$ it is clear that equalizing with one tap in the feedback path will only lower the BER
floor, whereas for 5 taps in the feedback path the error floor will be gone. Note that the BER
again only starts its sharp decrease at large $E_b/N_0$. The reason is of course the same as for
$\gamma = 7$. The strange result in Figure 3.39 is that the BER floor for one tap in the feedback path
is larger for $M_s = 4$ than for $M_s = 1$. This is not as expected and it could be that the MSE is
indeed lower, but that the BER is not. For 5 taps however the BER for $M_s = 4$ is again lower.
3.3 Decision feedback equalizer

Figure 3.38: BER curve for $\gamma = 7$. The BER for $\gamma = 7$ is lower than the BER for $M_s = 1$.

Finally we take a look at $\gamma = 10$. Here the dispersion is so high we do not remove the BER floor with our decision feedback filter. However we see that the error floor decreases for increasing $M_s$, as seen before, but also with increasing length of the feedback filter $h_{FB}$. In these figures the BER floor for $M_s = 4$ is always lower than the BER floor for the corresponding system with $M_s = 1$, as expected.

Figure 3.39: BER curve for $\gamma = 8$. 
3.3 Decision feedback equalizer

In all the previous simulations we use the original data symbols $a(k)$ in the feedback path. This is of course impossible in reality where we should apply the detected symbols to the feedback filter. Now we briefly investigate the difference between the two in terms of BER performance. To do so we have to simulate the BER, once for the feedback with the correct data symbols $a(k)$ and once for the feedback with the detected data symbols $\hat{a}(k)$. Because the decision of a symbol depends on previous decisions, we can no longer use the semi-analytical method and must simulate the BER in this section with a Monte-Carlo experiment.

For $\gamma = 4$ the plot of the BER can be found in Figure 3.41. Here the difference between the decision feedback with the estimated data symbols and the decision feedback with the correct data symbols is limited. When using the estimated data symbols, the BER is slightly larger than the BER for the feedback with $\hat{a}(k)$. For $\gamma = 7$ the influence of using $\hat{a}(k)$ is larger (Figure 3.38). Here the degradation at a BER of $1 \times 10^{-3}$ is approximately 3 dB.

To check why using the estimated symbols has more influence on $\gamma = 7$ than $\gamma = 4$, we take a look at the coefficients of the feedback filter with five taps at $E_b/N_0 = 15$ dB in Table 3.9. In this table we notice that the magnitudes of the coefficients for $\gamma = 7$ are larger than the magnitudes of the coefficients for $\gamma = 4$. This can be explained by the fact that for $\gamma = 7$, the added dispersion is larger and hence the NRZ pulse will be spread more by the channel. As a result there will be more ISI and we expect a filter that has more and larger meaningful terms. Due to fact that the coefficients of the feedback filter are larger, an error in the previous
estimates will have more influence on the current estimation. This will result in an increased probability of longer error events and hence a larger BER.

Table 3.9: Feedback filter coefficients at $E_b/N_0 = 15$ dB.

<table>
<thead>
<tr>
<th></th>
<th>$h_{FB}(1)$</th>
<th>$h_{FB}(2)$</th>
<th>$h_{FB}(3)$</th>
<th>$h_{FB}(4)$</th>
<th>$h_{FB}(5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 4$</td>
<td>0.31</td>
<td>0.043</td>
<td>0.0075</td>
<td>-0.0042</td>
<td>-0.0054</td>
</tr>
<tr>
<td>$\gamma = 7$</td>
<td>0.79</td>
<td>0.29</td>
<td>0.039</td>
<td>0.031</td>
<td>0.0090</td>
</tr>
</tbody>
</table>
3.3 Decision feedback equalizer

To support this reasoning we plot in Figure 3.43 the distribution of the length of the error events for $\gamma = 4$ at $E_b/N_0 = 15$. To obtain this distribution we simulate $1 \times 10^6$ data symbols. In the figure we see that for decision feedback with $a(k)$ we do not have error events longer than 1 error because at this value of $E_b/N_0$ the BER is already relatively low. For decision feedback with the estimated symbols $\hat{a}(k)$ we see that we now have some error events with length 2. This confirms the conclusion we derived from the BER curves: for $\gamma = 4$ using $\hat{a}(k)$ instead of $a(k)$ has not that much influence.

For $\gamma = 7$ the situation is different. We again plot in Figure 3.44 the distribution of the length of error events at $E_b/N_0 = 15$ dB with $1 \times 10^6$ data symbols. First we see that we have a lot more errors, which is of course a direct result of the higher BER. Next we observe that if we use the correct data symbols, we have a maximum of two errors in a row. For the decision feedback with the estimated symbols, the probability of longer error events increases. In our simulation this is especially visible for lengths three and four. This clearly shows that error propagation is a possible disadvantage of decision feedback.

(a) feedback with $a(k)$

(b) feedback with $\hat{a}(k)$

Figure 3.43: Length of error events for $\gamma = 4$.

(a) feedback with $a(k)$

(b) feedback with $\hat{a}(k)$

Figure 3.44: Length of error events for $\gamma = 7$. 
3.4 Non-linear decision feedback

3.4.1 Concept

Since the non-linearity in the PD generates quadratic terms of the data symbols \( a(k) \) it is not unthinkable that we can improve the equalizer by introducing linear combinations of those quadratic terms in the feedback path. The diagram is the same as for normal decision feedback, the only difference being that the feedback filter \( h_{FB} \) now becomes a Volterra filter.

3.4.2 Algorithm

Just as before, we minimize the MSE:

\[
\text{MSE} = \mathbb{E} \left[ |z(k) - a(k)|^2 \right]
\]

with

\[
z(k) = \sum_m h_{FF}(m) r((k)M_s + \tau - m) - \sum_{l>0} h_{FF}(l) \hat{a}(k-l)
\]

\[
- \sum_{Q} \sum_{p=1}^{P(q)} h'_{FB} \hat{a}(k-p) \hat{a}(k-p-q) - c_{off}
\]  

(3.15)

The calculation of the expression for the filter coefficients is again similar to the previous cases. To solve the system of linear equations we can again define a matrix \( R \) and vector \( r \). We performed this calculation once by hand, but because it is so elaborate and similar to the previous cases, we did not add it to the appendices.

As can be seen in equation 3.15, we have several degrees of freedom to decide which quadratic terms we include in the feedback. We have the parameter \( Q \) that determines how many periods there are maximally between the two data symbols. Then we have for each \( q \) a value \( P(q) \) that determines how many products we consider with \( q \) periods between the two data symbols. Due to the causality constraint (only decisions in the past are available) we can of course only include terms with past symbols. To make this more clear, we visualize the non-linear terms schematically in Figure 3.45 for an example where \( Q \) equals five and \( P = [2 3 2 4 2] \).
3.4.3 Simulation results

For the non-linear decision feedback we will first consider the case of $\gamma = 10$. The BER curves for this dispersion are given in Figure 3.46. We keep the forward filter at a length of 10 symbol periods and plot for both $M_s = 1$ and $M_s = 4$ the BER, once without decision feedback, once with the regular feedback and once with the non-linear feedback. On the figure we observe that linear feedback does not cancel the BER floor, whereas non-linear feedback does. In this figure the curves for $M_s = 4$ are situated below the corresponding curves for $M_s = 1$, as expected.

That the BER floor cannot be removed by an equalizer with these parameters for all dispersions can be seen in Figure 3.47. This figure is exactly the same as Figure 3.46 except that now $\gamma = 15$. 

Figure 3.46: BER curves in case of non-linear feedback and $\gamma = 10$.
3.5 Disadvantages of MSE criterion

In the figure we observe that for this dispersion the BER floor decreases, but is not gone. Further increasing the amount of linear and non-linear feedback taps and further increasing $M_s$ should however at one point lead to the disappearance of the BER floor.

![Figure 3.47: BER curves in case of non-linear feedback and $\gamma = 15$.]

$\gamma = 0$, no equalization
$\gamma = 15$, no equalization
$\gamma = 15$, $M_{\text{sample}} = 1$
$\gamma = 15$, $N_{\text{fb}} = 5$, $M_{\text{sample}} = 1$
$\gamma = 15$, $N_{\text{fb}} = 5$, $N_{\text{taps fb nl}} = [2 \ 2 \ 2 \ 2 \ 2]$, $M_{\text{sample}} = 1$
$\gamma = 15$, $M_{\text{sample}} = 4$
$\gamma = 15$, $N_{\text{fb}} = 5$, $M_{\text{sample}} = 4$
$\gamma = 15$, $N_{\text{fb}} = 5$, $N_{\text{taps fb nl}} = [2 \ 2 \ 2 \ 2 \ 2]$, $M_{\text{sample}} = 4$

3.5 Disadvantages of MSE criterion

In our algorithm we minimize the $\text{MSE} = E[|u(k) - a(k)|^2]$. This has as advantage that we get a linear system of equations that we can easily establish and solve to the equalizer coefficients. The disadvantage is that we actually want to minimize the BER, the next example shows that a minimum MSE does definitely not guarantee a minimum BER.

As in the simple example in section 3.2 we consider a system with an alternative channel pulse $p_{alt}(t)$ and a low-pass filter at $\frac{f_s}{2}$ instead of the TIA. The channel pulse is given by

$$
p_{alt}(t) = \begin{cases} \cos(\frac{\pi}{4} \frac{t}{T}) & \text{if } 0 \leq t \leq 2T \\ 0 & \text{elsewhere} \end{cases}
$$

(3.16)

We get ISI due to the $\cos(\frac{\pi}{4})$ at $t = T$, Figure 3.48 shows a histogram of $e(k)$ (as defined in equation 3.12) for $a(k) = 1$ and $a(k) = 0$ separately if we do not equalize. It is clear that this ISI will not lead to a BER floor.

Now we apply MMSE equalization to this system with $M_s = 1$ and 10 taps at either side. Figure 3.49 shows that the MSE has indeed decreased: the mass of $e(k)$ in the histogram is
3.5 Disadvantages of MSE criterion

![Histograms magnitude error unequalized system with \( p_{alt}(t) \).](image1)

Figure 3.48: Histograms magnitude error unequalized system with \( p_{alt}(t) \).

much more concentrated around 0. Our equalizer has however also introduced a BER floor by moving some of the histogram mass to \( e(k) > 0.5 \). This also becomes clear in the BER curves in Figure 3.50.

![Histograms magnitude error equalized system with \( p_{alt}(t) \).](image2)

Figure 3.49: Histograms magnitude error equalized system with \( p_{alt}(t) \).

Note that it is tempting to design our filters to minimize \( E[e(k)] \) instead of the MSE = \( E[e(k)^2] \). Because of the square in the MSE expression, negative \( e(k) \) will be corrected although they do not lead to a higher BER. If we differentiate the \( E[e(k)] \) expression to the filter coefficients, like we do with the MMSE equalizers, it is obvious that we do not get a system of equations because the filter coefficients in the \( E[e(k)] \) expression are not squared or multiplied with each other.
Figure 3.50: BER curves of system with $p_{alt}(t)$. 
Chapter 4

Continuous time equalizer

All the equalizers we have described so far operate on samples of the received signal. Equalizers are however often implemented as continuous analog devices as there are several benefits associated with equalization in the continuous-time domain. First of all, clock recovery only has to be performed on the equalized signal which is especially important if the eye is closed prior to equalization. Moreover, continuous time equalization will make the task of the clock recovery module easier. A next benefit is that the analog to digital converter (ADC) quantization noise will not be amplified by the equalizer. Lastly, the power consumption and chip area of the continuous time equalizer at a given data rate are generally lower than for its digital counterpart [21].

In this chapter we will adapt our algorithm for use with continuous-time linear feed forward equalizer. To verify the performance of our algorithm, we can perform measurements on a high-speed optical communication test setup equivalent to a PON.

4.1 Test setup

A schematic overview of the test setup is shown in Figure 4.1. It can be divided into three parts: transmitter, channel and receiver. In this section we describe this setup and pay particular attention to the characteristics of the equalizer.
4.1 Test setup

Figure 4.1: Test setup.

4.1.1 Transmitter

The carrier for our NRZ signal is provided by a 1547.8 nm CW laser followed by an erbium doped fiber amplifier (EDFA), isolator, filter, attenuator and polarizer. The filter is needed because an EDFA has a wavelength-dependent gain and leads to a broadened spectrum. We use a programmable attenuator to set the transmitter output power. The polarizer is needed to obtain the correct polarization for use with the MZM.

A 20 Gbit/s bit pattern is generated by the SHF 12100 B 50 Gbps bit pattern generator [22], whose clock is provided by the Anritsu MG3690B RF/Microwave signal generator [23]. During the measurements we will use a pseudorandom binary sequence (PRBS) pattern of length $2^{31} - 1$ to determine the BER and a known pattern of length $2^9$ to determine the equalizer coefficients (a PRBS pattern of length $2^9 - 1$ where we added a 1). These bit patterns are modulated on the carrier using the Fujitsu FTM7937EZ MZM [24].

4.1.2 Channel

We emulate the dispersion and attenuation of the channel with the TeraXion Clearspcetrum™ TDCMB benchtop tunable dispersion compensation module [25] and another attenuator. The dispersion module is used to compensate or emulate the dispersion caused by a long segment of glass optical fiber. Its dispersion parameter $DL$ can be set between -1200 ps/nm and 600 ps/nm; for our measurements we only use the positive settings because this corresponds to the dispersion of a normal optical fiber (negative dispersion fibers can be made and are used for dispersion compensation [9]). Because the chromatic dispersion parameter D(1550 nm) equals 15 ps/(km.nm) we can emulate fiber lengths between 0 and 40 km. At 20 Gbit/s the $DL$ settings...
correspond to $\gamma \approx DL/100$. Attenuator 2 sets the input power at the receiver side ($P_{PD}$), we investigate input powers between -20 and -12 dBm during our tests.

### 4.1.3 Receiver

At the receiver we perform the optical-electrical conversion with the Agilent 11982A Amplified Lightwave Converter [26]. It consists of a PIN PD and a low noise TIA. The electrical bandwidth of the TIA is only 11 GHz, but this is just right for our 20 Gbit/s data rate because we found in chapter 1 that the optimal TIA cutoff frequency $\approx 1/2T$. The 24 dB amplifier is included to get an electrical signal with peak-to-peak amplitudes in the order of 100 mV.

The SHF 11100B error analyzer [27] performs sampling, decision and comparison with the original bit pattern. It counts the number of bit errors and outputs an estimate for the BER. For every dispersion and attenuation setting we always first determine the timing and voltage offset of the detector which leads to minimal BER. First the analyzer minimizes the BER by performing a coarse automatic search then we tweak manually afterwards. If dispersion and/or noise close the eye too much, the error analyzer cannot sync to the signal and no BER can be determined. This is for example the case for a dispersion of 600 ps/nm, this is why it is not included in our measurements. When the eye is too open, on the other hand, no reliable BER can be estimated in a reasonable time span because too few bit errors occur. These two effects explain why some measurements are left blank.

The Agilent 86100C equivalent-time sampling oscilloscope [28] is used to measure eye diagrams and to sample the 2⁹ sequence. An equivalent-time sampling oscilloscope achieves high sample rates $1/T_{s,equivalent}$ by sampling a signal iteratively at a lower sample rate $1/T_{s,real}$: in a first cycle the signal is sampled at time instants $kT_{s,real}$, in cycle $n \in [0,T_{s,real}/T_{s,equivalent} - 1]$ we sample at $kT_{s,real} + nT_{s,equivalent}$. As a result we need to trigger the oscilloscope at 20/512 GHz (i.e. at the beginning of the PRBS sequence) to sample the 2⁹ sequence at a high equivalent sample rate.

### 4.1.4 Equalizer

The equalization in our test setup is performed by the Hittite HMC6545LP5E [29] continuous-time equalizer that can be bypassed to perform reference measurements. It consists of an input
4.1 Test setup

automatic gain control (AGC), a 9 tap delay line equalizer and an output driver. The AGC and output driver levels are set at an intermediate value and kept fixed during the measurements. The 9 equalizer taps are spaced at a fixed $T_{tap} = 18$ ps and their value can be programmed with a value between -63 and +63. The delay line consists of 8 delay elements where each element has 18 ps nominal propagation, the delayed signals are multiplied by the programmable tap amplifiers and then summed together.

Because of the analog implementation of the delay line and the multiplications we have to deal with some non-idealities. Figure 4.2 shows the frequency response of every equalizer tap. First of all it is clear that every tap has its own frequency response. We can see that the high frequency attenuation increases with tap number, this is a result of the accumulation of the low-pass characteristic of every delay element. We also see that the frequency response is definitely not flat over our frequencies of interest, setting one tap to its maximum gain and the others to zero will not pass the signal unaltered. On the contrary, the ISI will increase, resulting in an increased BER. Note that the relationship between the coefficients and small signal gain is not the same for each tap. Hence if we compensate for this non-ideality we must compensate each tap differently.

![Figure 4.2: Frequency response per tap, measured at coefficient value 63.](image-url)
The multiplication in the equalizer taps also introduces non-linear contributions to the signal. This non-linearity depends on the gain setting of the tap: higher gains introduce more non-linearity to the signal [29].

A last non-ideality is the relation between the programmed gain (an integer in the range of $[-63, 63]$) and the real small signal gain of the tap amplifier. We have a set of measurements at our disposal of the frequency responses of every tap at coefficient values -63, -32, 0, 32 and 63. Starting from these we determine the corresponding impulse responses and determine the gain as the amplitude of the highest peak in the impulse response. We interpolate these results and combine them in Figure 4.3.

![Figure 4.3: Fitted relationship between coefficient and small signal gain for every tap.](image)

### 4.2 Algorithm

It is clear from the previous discussion that we need to adapt our algorithm for use with a continuous time equalizer. First of all we neglect some non-idealities to be able to calculate the coefficients at all using a straightforward algorithm. We assume that the tap amplifiers do not
add non-linear contributions to the signal and that all tap frequency responses are equal.

The equalizer cannot add an offset $c_{off}$ to the signal but this no problem because the error analyzer used during the tests searches for the optimal offset itself. In our algorithm we do include $c_{off}$ factor in the MSE expression but simply do not use its calculated value.

We can now start deriving our algorithm, again starting from the MSE expression:

$$\text{MSE} = E \left[ |(v \star h_{eq})(kT + \tau) - c_{off} - a(k)|^2 \right]$$

$$(4.1)$$

With $T$ the symbol period, $T_{tap} = 18 \text{ ps}$ and $q = v \star h_{tap}$. Differentiating to $c_{off}$ and all $h(i)$ again leads to a system of equations: $Rh = r \Rightarrow h = rR^{-1}$ with $R_{(j,i)} = E[\tilde{q}(k, \tau, j)\tilde{q}(k, \tau, i)] - E[\tilde{q}(k, \tau, j)]E[\tilde{q}(k, \tau, i)]$ and $r_{(j)} = E[\tilde{q}(k, \tau, j)a(k)] - E[\tilde{q}(k, \tau, j)]E[a(k)]$ where $\tilde{q}(k, \tau, i) = q(kT + \tau - iT_{tap})$.

From these expressions we see that we need a measurement of $q(t)$ resulting from a known data symbol sequence $a$. The sample rate of this measurement should be high enough to represent $kT - iT_{tap}$ and to allow us to precisely choose the $\tau$-value that yields the equalizer with the lowest MSE. If we sample at a rate that is high enough to approximately fulfill the Nyquist-Shannon sampling theorem, this requirement can always be fulfilled by upsampling the signal.

Our algorithm now consists of 6 steps:

Step 1  Apply a known symbol sequence $a(k)$ to the system, e.g. the $2^9$ sequence we discussed in our description of the test setup. Set one tap of the equalizer to its maximal value and the others to zero. Sample the signal after the equalizer to get $q$. Set $\tau$ to an initial value.

Step 2  Determine the expected values $E[\tilde{q}(k, \tau, j)]$, $E[\tilde{q}(k, \tau, j)\tilde{q}(k, \tau, i)]$ and $E[\tilde{q}(k, \tau, j)a(k)]$ for the current value of $\tau$. Now we can fill the matrices $R$ and $r$.

Step 3  Calculate the coefficients $h$ corresponding to the system of equations $h = rR^{-1}$.

Step 4  Scale, adjust and quantize them so they assume values in the range $[-63, 63]$. In the adjustment step we compensate for the non-ideal relation between the programmed tap gain and actual small signal gain from Figure 4.3. We get $h_{programmable}$. 
Step 5 Readjust and scale $h_{\text{programmable}}$ back to $h_{\text{quantized}}$ to calculate the corresponding MSE = $\mathbb{E}[(\tilde{q}(k, \tau, 0) - C_{\text{off}} - a(k))^2]$. Note that we cannot use the expression $1/4 - r^T h_{\text{quantized}}$ because $h_{\text{quantized}} \neq h$!

Step 6 Carry out steps 2-5 for a range of $\tau$-values and select the result $(\tau, h_{\text{programmable}})$ that yields the lowest MSE. The selected $h_{\text{programmable}}$ are the MMSE filter coefficients. Note that we do not know the delay introduced by the system and measurement, as a result the range of $\tau$-values has to be large enough. Also note that $\tau$ defines the shift between the $\tilde{q}(k, \tau, 0)$ and $a(k)$ sequence and hence the location of the main equalizer tap.

To get consistent results, it is advisable to smooth the MSE curve as function of $\tau$ in step 6 of the algorithm, something we only discovered after carrying out our measurement. During our measurements, rerunning the algorithm with a set of samples taken with the same system some time later resulted in different optimal coefficients. The pragmatic approach taken during the measurements was to determine multiple sets of coefficients and choosing the one which leads to the best BER performance.

Figure 4.4 compares the MSE of the optimal equalizer coefficients at every $\tau$ for two measurements of the same system taken at a different time. We see that the general shape of the MSE curves is the same but that the small features in the detail of Figure 4.4 are different. From this figure it is clear that our algorithm will select a different $\tau$ for every measurement, corresponding to two sets of filter coefficients with a different location of the main tap.

The origin of these small differences is the limited length of the $a(k)$ sequence we used, this limits the accuracy of the MSE estimation. We could send a longer $a(k)$ sequence to overcome this problem, but this would increase the memory requirements and calculation time of our algorithm. An alternative way to solve this problem is smoothing the MSE curve in step 6 of our algorithm using a moving average filter. The curves in Figure 4.5 are now almost equal and the resulting filter coefficients almost match so this is clearly the preferred way to handle this in the future. Our ad hoc solution during the measurements was not as good as this one, but by selecting the best filter from multiple measurements we have at least approached the correct solution.

Before moving on, we take a closer look at the general shape of the MSE curves. We see that for $\tau$ in the approximate range $[-40, 80]$ (corresponding to $\approx 6T_{\text{tap}}$), the MSE is clearly lower than
4.2 Algorithm

Figure 4.4: Comparison MSE: measurement number 1 and 2, resulting coefficients:

Measurement 1: [16 -13 -23 -11 44 63 46 -36 -37]
Measurement 2: [-2 -20 7 63 30 -7 -34 -27 17]

Figure 4.5: Comparison smoothed MSE: measurement number 1 and 2, resulting coefficients:

Measurement 1: [16 -13 -23 -11 44 63 46 -36 -37]
Measurement 2: [14 -5 -33 -6 44 63 47 -43 -31]

for other $\tau$. For $\tau$ somewhere in the range of $[100, 140]$ we see a region of intermediate MSE values. Elsewhere we see a constant MSE of approximately 0.25. This is to be expected, at those $\tau$ values the equalized signal $r_{eq}(k)$ has no relation to $a_k$ and hence only half the equalizer output samples will be correct by chance, leading to an MSE of $0.25 = |0.5|^2$. 
4.3 Results

4.3.1 BER

At each dispersion setting we measure BERs for $P_{PD}$ ranging from -12 to -20 dBm. This once for the reference system where the equalizer is bypassed, once for the equalized system using the best coefficients found using our algorithm and once for the equalized system using the best coefficients found for a dispersion of 500 ps/nm. The equalizer coefficients are always determined once per dispersion setting at a $P_{PD}$ of -15 dBm. The resulting BER plots per dispersion setting are included in appendix D.2. In Figures 4.6-4.8 we give a summary.

![Figure 4.6: BER of reference system.](image-url)
4.3 Results

Figure 4.7: BER of equalized system.

Figure 4.8: BER of system equalized with coefficients of system 500 ps/nm.
Unexpected result

In Figure 4.9 we also show the BER curves at a dispersion of 300 ps/nm because here we see something unexpected: the BER of the system equalized with the coefficients determined at 500 ps/nm is lower than when equalized with the coefficients determined at 300 ps/nm!

To make sure that our algorithm is not wrong, we plot in Figure 4.10 two MSE curves. $MSE_{300}$ shows the MSE of the optimal equalizer coefficients for every $\tau$, $MSE_{500}$ shows the MSE corresponding to the optimal coefficients determined at 500 ps/nm, maximized over $\tau$. Note that the latter corresponds with one fixed $\tau$. This explains the shape of the MSE curve: we only have a low MSE over the limited range of $\tau$-values close to the original optimal $\tau$. $MSE_{500}$ is never lower than $MSE_{300}$ so the difference in BER has another cause.

A possible explanation is the influence of the equalizer AGC together with the scaling of the coefficients in step 4 of our algorithm to the range $[-63, 63]$ (the original coefficients minimize the MSE between the equalized signal and the original data symbols $\in \{0, 1\}$). To get the best resolution we do this by scaling each set of found coefficients so they each have their maximum coefficient at $-63$ or $+63$, as a result the scaling factor differs between different sets of coefficients.
We can do this because the error detector scales its input signal back to the correct range by searching for a scaling factor that minimizes the BER.

To see why this leads to the unexpected result, we present in Figure 4.11 a schematic overview of the relevant receiver components. At the input of the equalizer we receive $s(t) + n_1(t)$ where $s(t)$ is the signal contribution at the TIA output and $n_1(t)$ the noise added before the equalizer. At the error analyzer we receive

$$\alpha[h * (s(t) + n_1(t))] + n_2(t) = (\beta_{\text{coef}} h) * [\beta_{\text{AGC}}(s(t) + n_1(t))] + n_2(t)$$

(4.2)

The impulse response of the equalizer is denoted by $h$. The factor $\alpha$ combines the scaling of the equalizer coefficients $\beta_{\text{coef}}$ and the influence of the AGC $\beta_{\text{AGC}}$; $n_2(t)$ is the noise added after the equalizer (e.g. in the output driver of the equalizer or in the input circuit of the error analyzer).

The signal to noise ratio (SNR) at the equalizer input then equals $P_s/P_{n_1}$ whereas at the input of the error analyzer we get

$$SNR_{\text{dec}} = \frac{\alpha^2 P_s}{\alpha^2 P_{n_1} + P_{n_2}} = \frac{P_s}{P_{n_1} + \frac{P_{n_2}}{\alpha^2}}$$

(4.3)

Here we use the notation $P_x$ for the power of $x$ and $P_{\tilde{x}}$ for the power of $\tilde{x} = h * x$ where $x = s$, $\tilde{x} = s * h$.
4.3 Results

Figure 4.11: Schematic overview receiver.

\[ s(t) + n_1(t) \rightarrow \text{Equalizer} \rightarrow SNR_{\text{dec}} \rightarrow \text{Decision} \]

\[ \beta_{\text{AGC}} \]

\[ \beta_{\text{eq}} \]

\[ n_2(t) \]

\[ n_1 \text{ or } n_2. \text{ We also assume that } P_{n_1} > P_{n_2}, \text{ i.e. the main noise contributors are located before the equalizer.} \]

Now assume that one set of coefficients is scaled with \( \beta_{\text{coeff},1} \) and another with \( \beta_{\text{eq},2} > \beta_{\text{eq},1} \). For low \( P_{PD} \), \( \beta_{\text{AGC},1} = \beta_{\text{AGC},2} \) will be high resulting in \( SNR_{\text{dec},2} \approx SNR_{\text{dec},1} \) because \( \alpha^2 P_{n_1} \gg P_{n_2} \).

If we however increase \( P_{PD} \), \( \beta_{\text{AGC},1} = \beta_{\text{AGC},2} \) will decrease resulting in a \( P_{n_2} \) that is no longer negligible. This leads to \( SNR_{\text{dec},2} > SNR_{\text{dec},1} \). That is exactly what we see in Figure 4.9: the difference between the BER curves increases for increasing \( P_{PD} \).

If we check the scaling factors of the different sets of equalizer coefficients we determined, our suggested explanation seems to be true. The scaling factor of the coefficients at 500 ps/nm is 0.56, at 300 ps/nm we have a factor of 0.47. For the other sets of coefficients we find that they are all greater than 0.54 except for the coefficients at 400 ps/nm where we have a factor 0.45. Figure D.11 in Appendix D.2 shows that the coefficients determined at a dispersion of 500 ps/nm also perform best at 400 ps/nm so this reinforces our theory.

Note that a factor that of course also plays a role in the BER results is the overcompensation of the dispersion resulting from using the 500 ps/nm coefficients in situations with less dispersion. The negative effect of this at 300 and 400 ps/nm seems to be undone by the difference in scaling factor.

Power gain

Instead of expressing the influence of the equalizer in terms of BER improvement, we can also determine a power gain: how much dB the \( P_{PD} \) of the reference system has to increase in order to get the same BER performance as the equalized system. This power gain is shown in Figures 4.12-4.13, again once for the regular coefficients and once for the coefficients of 500 ps/nm dispersion.

We see that the gain of the equalizer depends on two factors: the dispersion and the BER.
The former is of course because more gain can be achieved if dispersion deteriorated the signal more. The latter is because the distance between the equalized and non-equalized BER curves increases for decreasing BER. This we also see in the BER plots of the discrete time equalizers, especially when a BER floor is present.

Now we can interpret these power gains in terms of parameters of the PON network. First of all, the equalizer power gain can be used to reduce the transmitted power. This results in a lower power consumption at the OLT and ONU. Of course we also have to take the extra power consumption of the equalizer into account, \( \approx 400 \text{ mW or 26 dBm} \) according to the datasheet [29]. Placing an equalizer at the OLT leads to the largest gains as the transmit power of all ONUs can be decreased. Note that an equalizer at the OLT cannot be optimal for all ONUs but we expect that optimizing for the furthest ONU or an ONU at an intermediate distance will be beneficial for all ONU. This can be seen in the performance of the equalizer coefficients determined at 500 ps/nm at other dispersion settings (cf. Figure 4.13).

We can also use the power gain to increase the split ratio, i.e. increase the number of served
4.3 Results

Figure 4.13: Power gain of system equalized with coefficients of system 500 ps/nm.

customers. If we ignore the excess splitter losses, we get:

\[ N_{\text{users,new}} = \left\lfloor 10^{\text{Power gain}/10} N_{\text{users,old}} \right\rfloor \]  

(4.4)

Finally we can also increase the maximum fiber length. If we use the fiber attenuation of \( \alpha = 0.5 \text{ dB/km} \) from Chapter 1, we get:

\[ L_{\text{new}} = \alpha \text{ Power gain} + L_{\text{old}} = 0.5 \text{ Power gain} + L_{\text{old}} \]  

(4.5)

But increasing the fiber length also increases the dispersion and as a result translating the power gain to an increased maximum fiber length is not straightforward. Using the Figures in appendix D.2 we can for example investigate going from a dispersion of 200 ps/nm to 300 ps/nm at a BER of \( 10^{-6} \). With the reference system at 200 ps/nm we need \( P_{PD} = -13 \text{ dBm} \) to get a BER of \( 10^{-6} \), in the equalized system at 300 ps/nm we need \( P_{PD} = -15.5 \text{ dBm} \). The 2.5 dB power gain is however not sufficient to undo the extra attenuation of

\[ \frac{0.5 \text{ dB}}{15 \text{ ps/nm.km}} \times 100 \text{ ps/nm} = 0.5 \frac{\text{dB}}{\text{km}} \times 6.67 \text{ km} = 3.35 \text{ dB} \]  

(4.6)

We remind that we assume that \( D = 15 \frac{\text{ps}}{\text{nm.km}} \). The power at the transmitter hence still needs to be increased by 0.85 dB.
4.3 Results

4.3.2 Eye diagrams

In appendix D.1 we present the eye diagrams corresponding to all BER measurements we performed. Comparing the eye diagrams of the reference system (no equalization) and the equalized system is not straightforward because of the equalizer AGC. As a result, the eye openings before and after equalization cannot be compared. The shape, however, can be compared and in the case of 500 ps/nm dispersion we can clearly see the equalizer opening the eye. These eye diagrams are shown in Figure 4.14.

![Eye diagrams](image)

(a) Reference system  
(b) Equalized system

Figure 4.14: Eye diagrams at a dispersion of 500 ps/nm.

4.3.3 Equalizer responses

In Figure 4.15 we present the impulse response and frequency response of the filters determined by our algorithm at 100 ps/nm and 500 ps/nm. In Figure 4.15(a) we see that the maximum of both filters occurs at a different time instant, this is because the filters do not have the same $\tau$-value. We also see that the 500 ps/nm filter deviates more from the Dirac-impulse than the 100 ps/nm filter because it obviously has to compensate more ISI. Figure 4.15(b) shows the frequency responses, here we cannot tell much except that the maximum for both is situated at approximately 10 GHz.
4.3 Results

Figure 4.15: Comparison equalizers at 100 ps/nm and 500 ps/nm.

(a) Impulse response

(b) Frequency response
Chapter 5

Conclusions

5.1 Conclusions

To conclude this thesis, we end with an overview of our work and what we achieved. Opportunities for future work are discussed in the section hereafter.

First we have developed a model of a PON communication system where only the most important components are included. This allowed us to condense the PON parameters into a minimal number of model parameters, simplifying setting up simulations. To model the optical fiber chromatic dispersion, we derived an analytical expression for the fiber impulse response $h_{fib}$. As this impulse response has an infinite length and does not extinguish, we had to consider $p_{fib}$, the response of the fiber to an NRZ pulse, instead. This $p_{fib}$ is also infinite but does extinguish allowing us to truncate it at a certain length. Care has to be taken when doing this as the length at which we truncate $p_{fib}$ is limited by the simulation sample rate.

If we consider a system without chromatic dispersion, the PD non-linearity has no influence and an analytical treatment is possible. We derived an expression for the BER in function of system parameters and used it to optimize the TIA. This trade-off between noise bandwidth maximal ISI yielded an underdamped TIA with cutoff frequency at approximately half the symbol rate.

If we introduce dispersion the PD non-linearity now does have an influence. An analytical treatment was no longer possible so we had to simulate the eye-diagram without noise to get an idea of the worst-case BER. Increasing the dispersion parameter $\gamma$ gradually closes the eye until
5.1 Conclusions

it is completely closed at $\gamma \approx 6$.

To get an idea of the performance of our system we had to implement a BER simulation. For this we used a semi-analytical approach as the Monte-Carlo simulations proved to be too time consuming.

Next we applied discrete time equalization to our system in order to improve its performance. Starting from the MMSE criterion, a system of equations in the equalizer parameters was derived. A first equalizer we investigated was the linear equalizer, i.e. its output is a linear combination of the received signal sampled at the symbol rate. We found that the maximal eye opening is no longer the optimal sample moment when equalizing, instead we go over a range of sample moments around the maximal eye opening and select the one leading to the minimal MSE after equalization.

We simulated the BER performance of this linear equalizer at low, medium and high dispersion. At low dispersion, equalizing has almost no influence because there is not much ISI to compensate. For a medium amount of dispersion the equalizer manages to decrease the BER significantly at high SNR. At high dispersion the eye is closed before equalization, resulting in a BER floor. Equalization with the linear equalization only lowers this BER floor for dispersion $\gamma > 6$. So this equalizer is very straightforward to implement and calculate, but it has a poor performance in a system with high dispersion. The reason for this poor performance is the PD non-linearity, we verified this by simulating our system where we replaced the PD by a linear component. In that case equalization removed the BER floor.

To solve our issues with the BER floor, we considered a fractionally spaced linear amplifier. Its output still consists of a linear combination of samples of the received signal, but this time sampled at a multiple of the symbol rate $1/T$. Conceptually we can still use the same algorithm to determine the filter parameters, we only had to adapt the expression of the linear system of equation. We expected this equalizer to outperform our first equalizer because its increased degrees of freedom allow it to better handle the cross terms in the received signal as a result of the PD non-linearity. We derived an expression that tells us what the minimal oversampling rate $M_s$ will be for given channel filter length to obtain a ZF equalizer. To verify this, we first studied a simple system where $p_{fib}$ is replaced by an alternative, short pulse. We made use of histograms of the error $e_k$ to interpret the BER results, the existence of a BER floor is also easily seen on these histograms. For this simple system it was possible to remove the BER floor at the expected oversampling $M_s$ of the received signal.
We again simulated BER performance at low, medium and high dispersion. At a low and medium dispersion, the performance improvement as a result of increasing $M_s$ is negligible. For a high value of the dispersion parameter, the BER floor moves lower for increasing $M_s$. Although we expect that by increasing $M_s$ we could completely remove the BER floor.

To further improve the performance we also studied a decision feedback equalizer. Here we also use past symbol estimates, as a result our equalizer is no longer linear. After adapting the expression for the linear system of equations, we could again use the same algorithm to determine our filter coefficients. Because the output of this equalizer depends on past decisions, we could no longer use the semi analytical approach to simulate the BER. To circumvent this limitation, we used the correct symbols $a_k$ instead of the estimated $\hat{a}_k$. A Monte-Carlo simulation of the real decision feedback equalizer showed that the difference in BER performance is small. Here the $\gamma = 7$ BER floor already disappears for a filter with $M_s = 1$ and only one feedback tap! For larger dispersion decision feedback will only lower the BER floor.

As non-linear terms are created by the PD it is expected that non-linear feedback will improve the systems performance. In the MSE expression an extra non-linear term appears, but we can still use the same algorithm. The simulation showed that the non-linear equalizer outperforms the other equalizers, this was above all visible for large dispersion, where the non-linear equalizer solves the BER floor where the other equalizer could not at a comparable complexity.

All equalizers we studied minimize the MSE at the decision unit. A minimal MSE however does not lead to the minimal BER we are actually interested in. To demonstrate this we gave a simple example of a system where equalization increases the BER and even introduces a BER floor.

We also extended our scope to continuous time equalizers. These obviously behave differently from discrete time equalizer and suffer from some non-idealities, but by applying some suitable approximations we were able to again adapt the expression for our linear system of equations and formulated a practical algorithm for use with the PON emulation test setup. To get consistent results from our algorithm, we had to add a smoothing operation in the last step of our algorithm. We only discovered this after performing our measurements, so we had to use an ad hoc solution during the measurements.

The measurements confirmed that our algorithm works as it increased the system performance. We achieved power gains from approximately 1 to 5 dB depending on the BER and dispersion
of the test setup. During the measurements we also got some anomalous results where the coefficients determined at one dispersion setting outperformed those at the other dispersion setting the system was at. We offered an explanation of this based on our algorithm and the test setup details. This explanation is also supported by what we see on the measured BER curves.

5.2 Future work

One subject that further can be explored is the modulation scheme. In this thesis we always used a NRZ pulse to modulate the data. The return-to-zero (RZ) pulse could be an alternative. This pulse is only high for a fraction of the bit period. Since these pulses are shorter in time, they can tolerate more spreading and hence more dispersion. A disadvantage is that this pulse occupies a larger bandwidth.

Another possibility is Duobinary Modulation. The big advantage of this scheme is that it uses less than R/2 Hz for transmitting R bits/sec. By performing some operations we create ISI at the transmitter that result in longer pulses in the time domain. The ISI is applied in such a way that it can be cancelled at the receiver and the data can be perfectly detected. Due to the longer pulses in time domain, the spectrum is smaller and hence the distortion effects of the channel will be smaller [30].

To improve the data rate while occupying the same bandwidth, we could consider more complex constellations. In this thesis we limit the data symbols to the set of \( \{0, 1\} \), a two-level constellation. Using a higher order constellation (e.g. M levels) we can map \( \log_2(M) \) bits to one symbol, resulting in a \( \log_2(M) \) larger data rate. The disadvantage is that the noise margin will decrease, which results in an increased probability of detecting an erroneous symbol.

To calculate the equalizer coefficients we send a known sequence \( \mathbf{a} \) of data symbols through the system and try to minimize the MMSE at the output. The disadvantage is that if the systems changes, in particular the fiber characteristics, the coefficients are no longer optimal. A possible solution is adapting the coefficients on a regular basis. The temperature for example, which often changes over time, will have an influence on the dispersion parameter [31].
Bibliography


[22] SHF 12100 B 50 Gbps Bit Pattern Generator datasheet, SHF Communication Technologies AG, Berlin, Germany, 2009

[23] MG3690B RF/Microwave Signal Generator datasheet, Anritsu Corporation, Japan, 2011

[24] Low Drive Voltage 40 Gbps LiNbO$_3$ External Modulator datasheet, Fujitsu Ltd, Japan, 2009


[26] 11982A Amplified Lightwave Converter datasheet, Agilent Technologies, USA, 2004
[27] SHF 11100 B Error Analyzer datasheet, SHF Communication Technologies AG, Berlin, Germany, 2009


[29] HMC6545LP5E 32 Gbps Dual Channel Advanced Linear Equalizer datasheet, Hittite Microwave Corporation, Chelmsfort, USA, 2013


Appendices
Appendix A

Fiber impulse response

A.1 Calculation $h_{fib}$

$$h_{fib}(u) = \int_{-\infty}^{+\infty} H_{fib}(f) e^{j2\pi fu} df$$

$$= \int_{-\infty}^{+\infty} e^{-j\gamma f^2 T^2} e^{j2\pi fu} df$$

$$= \int_{-\infty}^{+\infty} e^{-a(f-\frac{b}{2a})^2 + \frac{b^2}{4a}} df$$

with $a = j\gamma T^2$ and $b = j2\pi u$

$$= e^{\frac{b^2}{4a}} \int_{-\infty}^{+\infty} e^{-a(f-\frac{b}{2a})^2} df$$

$$= e^{\frac{b^2}{4a}} \sqrt{\frac{\pi}{a}} = \frac{1}{T} \sqrt{\frac{-j\pi}{\gamma}} e^{\frac{\gamma^2 u^2}{2T^2}}$$

$$= \sqrt{\frac{\gamma}{2T}} \sqrt{\frac{\pi}{\gamma}} \left( \cos \left( \frac{\pi^2 u^2}{\gamma T^2} \right) + \sin \left( \frac{\pi^2 u^2}{\gamma T^2} \right) \right) +$$

$$j \frac{\sqrt{2}}{2T} \sqrt{\frac{\pi}{\gamma}} \left( \sin \left( \frac{\pi^2 u^2}{\gamma T^2} \right) - \cos \left( \frac{\pi^2 u^2}{\gamma T^2} \right) \right)$$

(A.1)

To calculate the integral we use that

$$\int_{-\infty}^{+\infty} e^{-jp(x+c)^2} dx = \frac{\pi}{\sqrt{jp}} \text{ for } p \in \mathbb{R}, c \in \mathbb{C} \text{ and } p > 0$$

(A.2)

which we prove in appendix A.2.
A.2 Gaussian integral with complex offset

Theorem:
\[ \int_{-\infty}^{+\infty} e^{-jp(x+c)^2} \, dx = \sqrt{\frac{\pi}{jp}} \quad \text{for } p \in \mathbb{R}, \ c \in \mathbb{C} \text{ and } p > 0 \quad (A.3) \]

Proof based on [32]:

We write \( c = a + jb \) where we can assume \( a = 0 \) without loss of generality because this can always be achieved by performing the real substitution \( x = x - a \). Next we define \( f(z) = e^{-jpz^2} \) which is analytic over the whole complex plane. (Proof: Cauchy-Riemann equations valid over whole complex plane)

As a result of Cauchy’s integral theorem we then get that
\[ \int_{-T}^{T} f(x) \, dx + \int_{0}^{b} f(T + jy) \, jdy + \int_{T}^{-T} f(x + jb) \, dx + \int_{b}^{0} f(-T + jy) \, jdy = 0 \quad (A.4) \]

where \( T \in \mathbb{R} \), \( x \) and \( y \) are real variables and the paths of the four integrals constitute a contour in the complex plane.

For \( T \to +\infty \), (1) can be written as
\[
2 \left( \int_{0}^{+\infty} \cos(pz^2) \, dz - j \int_{0}^{+\infty} \sin(pz^2) \, dz \right) = \frac{2}{\sqrt{p}} \left( \int_{0}^{+\infty} \cos(x^2) \, dx - j \int_{0}^{+\infty} \sin(x^2) \, dx \right) = \sqrt{\frac{\pi}{jp}} \quad (A.5)
\]

where we used that the limits of the Fresnel integrals are \( \int_{0}^{+\infty} \cos(x^2) \, dx = \int_{0}^{+\infty} \sin(x^2) \, dx = \sqrt{\frac{\pi}{8}} \)

The sum of (2) and (4) goes to zero for \( T \to +\infty \) as for \( T \to +\infty \) we get that \( e^{-jp(T+jb)^2} = e^{-jp(-T+jb)^2} \to e^{-jpT^2} \) which leads to
\[
(2) + (4) = \int_{0}^{b} e^{-jpT^2} \, jdy + \int_{b}^{0} e^{-jpT^2} \, jdy = 0 \quad \text{for } T \to +\infty \quad (A.6)
\]

As a result we get that (1) is equal to \(-(3)\) which is what we wanted to prove.
Appendix B

Derivation expression for $s(t)$

\[ s(t) = \int h_{\text{tia}}(t-u) i_{\text{pd}}(u) \, du \]
\[ = \int h_{\text{tia}}(t-u) \left( \sum_{m} a(m) p_{\text{fib}}(u-mT) \right)^2 \, du \]
\[ = \int h_{\text{tia}}(t-u) \left( \sum_{m} a(m) p_{\text{fib}}(u-mT) \right) \left( \sum_{m'} a(m') p_{\text{fib}}(u-m'T) \right)^* \, du \]
\[ = \sum_{m} a(m) \int h_{\text{tia}}(t-u) p_{\text{fib}}(u-mT) \sum_{m'=m-M}^{m+M} a(m') p_{\text{fib}}^*(u-m'T) \, du \]
\[ = \sum_{m} a(m) \int h_{\text{tia}}(t-u) p_{\text{fib}}(u-mT) \sum_{k=-M}^{M} a(m+k) p_{\text{fib}}^*(u-kT) \, du \]
\[ = \sum_{m} a(m) \sum_{k=-M}^{M} a(m+k) \int h_{\text{tia}}(t-u-mT) p_{\text{fib}}^*(u-kT) \, du \]
\[ = \sum_{m} a(m) \sum_{k=-M}^{M} a(m+k) \int h_{\text{tia}}(t-u-mT) p_{\text{fib}}^{(2)}(u) \, du \]
\[ = \sum_{m} a(m) \sum_{k=-M}^{M} a(m+k) h_{\text{tot},k}(t-mT) \] (B.1)

with $p_{\text{fib}}(t) \approx 0$ for $t \notin [-MT, MT]$, $p_{\text{fib}}^{(2)}(t-kT) = p_{\text{fib}}(t)p_{\text{fib}}^*(t-kT)$ and $h_{\text{tot},k}(t) = \int h_{\text{tia}}(t-u)p_{\text{fib}}^2(u) \, du$. 
Appendix C

Calculation of equalizer coefficients

C.1 Calculation of expression MSE

In this section we give the derivation of the simple expression for the MSE for the linear equalizer:

\[
MSE = E \left[ \left| \sum_m h_{FF}(m) r(k-m) - c_{off} - a(k) \right|^2 \right]
\]

\[
= E \left[ \left| \sum_m h_{FF}(m) r(k-m) - \sum_m h_{FF} E[r(k-m)] + E[a(k)] - a(k) \right|^2 \right]
\]

\[
= E \left[ \left| \sum_m h_{FF}(m)(r(k-m) - E[r(k-m)]) - (a(k) - E[a(k)]) \right|^2 \right]
\]

\[
= E \left[ \sum_{m,m'} h_{FF}(m) h_{FF}(m')(r(k-m) - E[r(k-m)]) (r(k-m') - E[r(k-m')])
\]

\[
-2 \sum_m h_{FF}(m)(r(k-m) - E[r(k-m)]) (a(k) - E[a(k)])
\]

\[
+ (a(k) - E[a(k)])^2 \right]
\]

\[
= \sum_{m,m'} h_{FF}(m) h_{FF}(m')(E[r(k-m)r(k-m')] - E[r(k-m)] E[r(k-m')])
\]

\[
-2 \sum_m h_{FF}(m)(E[s(k-m)a(k)] - E[s(k-m)] E[a(k)])
\]

\[
+ E[|a(k)|^2] - E[a(k)]^2 \right]
\]

\[
= h_{FF}^T R h_{FF} - 2 h_{FF}^T r + \frac{1}{2} - \frac{1}{4}
\]

\[
= r^T R^{-1} R r - 2 r^T R^{-1} r + \frac{1}{4}
\]

\[
= \frac{1}{4} - r^T R^{-1} r
\]  \hspace{1cm} (C.1)
C.2 Linear equalization with multiple samples per symbol

We have to minimize the expression:

\[ \text{MSE} = E \left[ |u(k) - a(k)|^2 \right] = E \left[ \sum_m h_{FF}(m) r(kM_s + \tau - m) - c_{off} - a(k)|^2 \right] \quad (C.2) \]

To minimize expression C.2 we equate the derivatives with respect to the unknowns: \( h_{FF}(m) \) and \( c_{off} \) to zero. The derivatives are:

\[
\begin{align*}
\frac{\partial \text{MSE}}{\partial h_{FF}(m')} &= E \left[ 2r(kM_s + \tau - m') \left( \sum_m h_{FF}(m) r(kM_s + \tau - m) - a(k) - c_{off} \right) \right] \quad \forall m' \\
\frac{\partial \text{MSE}}{\partial c_{off}} &= E \left[ -2 \left( \sum_m h_{FF}(m) r(kM_s + \tau - m) - a(k) - c_{off} \right) \right]
\end{align*}
\]

Equating to zero leads to:

\[
\begin{align*}
0 &= \sum_m h_{FF}(m) E[r(kM_s + \tau - m) r(kM_s + \tau - m')] - E[a(k) r(kM_s + \tau - m')] \\
\Rightarrow 0 &= \sum_m h_{FF}(m) E[r(kM_s + \tau - m)] - E[a(k)] \\
\Rightarrow c_{off} &= \sum_m h_{FF}(m) E[r(kM_s + \tau - m)] - E[a(k)] \\
\Rightarrow 0 &= \sum_m h_{FF}(m) (E[r(kM_s + \tau - m') r(kM_s + \tau - m)] - E[r(kM_s + \tau - m')] E[r(kM_s + \tau - m)]) \\
&- (E[a(k) r(kM_s + \tau - m')] - E[a(k)] E[r(kM_s + \tau - m')]) \quad \forall m' \quad (C.3)
\end{align*}
\]

Define \( N_{FF} = N_{FF,L} + N_{FF,R} + 1 \) as the sum of the number of taps left \((N_{FF,L})\) and right \((N_{FF,R})\) in the forward filter. Define the \( N_{FF} \times N_{FF} \) matrix \( R \) with \( R(i,j) = E[r(kM_s + \tau - m') r(kM_s + \tau - m)] - E[r(kM_s + \tau - m')] E[r(kM_s + \tau - m)] \) with \( m = i - N_{FF,L} - 1 \) and \( m' = j - N_{FF,R} - 1 \). Further make the vector \( r \) with \( r(i) = E[a(k) r(kM_s + \tau - m)] - E[a(k)] E[r(kM_s + \tau - m)] \) again with \( m = i - N_{FF,L} - 1 \). Now we can rewrite equation C.3 in terms of matrices:

\[ R h_{FF} = r \rightarrow h_{FF} = r R^{-1} \]

In this equation the vector \( h_{FF} \) contains the filter coefficients.

C.3 Linear equalization with decision feedback

We have to minimize the expression:

\[ \text{MSE} = E \left[ |z(k) - a(k)|^2 \right] = E \left[ \sum_m h_{FF}(m) r(kM_s + \tau - m) - \sum_{l>0} h_{FB}(l) \hat{a}(k-l) - c_{off} - a(k)|^2 \right] \quad (C.4) \]
To minimize expression C.4 we equate the derivatives with respect to the unknowns: \( h_{FF}(m), h_{FB}(l) \) and \( c_{off} \) to zero. The derivatives are:

\[
\frac{\partial \text{MSE}}{\partial h_{FF}(m')} = E \left[ 2r(kM_s + \tau - m') (\sum_m h_{FF}(m) r(kM_s + \tau - m) - \sum_{l>0} h_{FB}(l) \hat{a}(k-l) - a(k) - c_{off}) \right] \forall m'
\]

\[
\frac{\partial \text{MSE}}{\partial h_{FB}(l')} = E \left[ -2 \hat{a}(k-l') (\sum_m h_{FF}(m) r(kM_s + \tau - m) - \sum_{l>0} h_{FB}(l) \hat{a}(k-l) - a(k) - c_{off}) \right] \forall l'
\]

\[
\frac{\partial \text{MSE}}{\partial c_{off}} = E \left[ -2 (\sum_m h_{FF}(m) r(kM_s + \tau - m) - \sum_{l>0} h_{FB}(l) \hat{a}(k-l) - a(k) - c_{off}) \right]
\]

Equating to zero leads to:

\[
0 = + \sum_m h_{FF}(m) E[r(kM_s + \tau - m') r(kM_s + \tau - m)] - \sum_{l>0} h_{FB}(l) E[\hat{a}(k-l) r(kM_s + \tau - m')]
- E[a(k) r(kM_s + \tau - m')] - c_{off} E[r(kM_s + \tau - m')] \forall m'
\]

\[
0 = + \sum_m h_{FF}(m) E[r(kM_s + \tau - m) \hat{a}(k-l')] - \sum_{l>0} h_{FB}(l) E[\hat{a}(k-l) \hat{a}(k-l')]
- E[a(k) \hat{a}(k-l')] - c_{off} E[\hat{a}(k-l')] \forall l'
\]

\[
c_{off} = \sum_m h_{FF}(m) E[r(kM_s + \tau - m)] - \sum_{l>0} h_{FB}(l) E[\hat{a}(k-l)] - E[a(k)]
\]

Define \( N_{FF} = N_{FF,L} + N_{FF,R} + 1 \) as the sum of the number of taps left(\( N_{FF,L} \)) and right(\( N_{FF,R} \)) in the forward filter and \( N_{FB} \), the number of taps in the feedback filter. Create the (\( N_{FF} + \))
C.3 Linear equalization with decision feedback

\[ N_{FB} \times (N_{FF} + N_{FB}) \] matrix \( R \):

\[
R(i, j) = \begin{cases}
    i \leq N_{FF} \& j \leq N_{FF} : & E[r(kM_s + \tau - n)r(kM_s + \tau - m)] - E[r(kM_s + \tau - n)]E[r(kM_s + \tau - m)] \\
    i \leq N_{FF} \& j > N_{FF} : & -(E[\hat{a}(k - q)r(kM_s + \tau - n)] - E[\hat{a}(k - q)]E[r(kM_s + \tau - n)]) \\
    i > N_{FF} \& j \leq N_{FF} : & E[\hat{a}(k - p)r(kM_s + \tau - m)] - E[\hat{a}(k - p)]E[r(kM_s + \tau - m)] \\
    i > N_{FF} \& j > N_{FF} : & -(E[\hat{a}(k - p)\hat{a}(k - q)] - E[\hat{a}(k - p)]E[\hat{a}(k - q)])
\end{cases}
\]

with \( n = i - N_{FF,L} - 1 \), \( m = j - N_{FF,L} - 1 \), \( p = i - N_{FF} \) and \( q = j - N_{FF} \). Next define the \((N_{FF} + N_{FB}) \times 1\) vector \( r \):

\[
r(i) = \begin{cases}
    i \leq N_{FF} : & E[a(k)r(kM_s + \tau - n)] - E[a(k)]E[r(kM_s + \tau - n)] \\
    i > N_{FF} : & E[a(k)\hat{a}(k - p)] - E[a(k)]E[\hat{a}(k - p)]
\end{cases}
\]  

(C.2)

Then the equalize coefficients \( h_{eq} = [h_{FF}; h_{FB}] \) can be calculated as:

\[
Rh_{eq} = r \Rightarrow h_{eq} = R^{-1}r
\]

(C.3)
Appendix D

Measurement results

D.1 Eye diagrams per dispersion setting

(a) Reference system

(b) Equalized system

Figure D.1: Eye diagrams at a dispersion of 0 ps/nm.
D.1 Eye diagrams per dispersion setting

Figure D.2: Eye diagrams at a dispersion of 100 ps/nm.

(a) Reference system

(b) Equalized system

Figure D.3: Eye diagrams at a dispersion of 200 ps/nm.

(a) Reference system

(b) Equalized system

Figure D.4: Eye diagrams at a dispersion of 300 ps/nm.

(a) Reference system

(b) Equalized system
D.1 Eye diagrams per dispersion setting

Figure D.5: Eye diagrams at a dispersion of 400 ps/nm.

(a) Reference system

(b) Equalized system

Figure D.6: Eye diagrams at a dispersion of 500 ps/nm.

(a) Reference system

(b) Equalized system
D.2 BER plots per dispersion setting

Figure D.7: BER at dispersion of 0 ps/nm.
Figure D.8: BER at dispersion of 100 ps/nm.

Figure D.9: BER at dispersion of 200 ps/nm.
Figure D.10: BER at dispersion of 300 ps/nm.

Figure D.11: BER at dispersion of 400 ps/nm.
Figure D.12: BER at dispersion of 500 ps/nm.