EXPERIMENTAL DEMONSTRATION OF MITIGATION OF LINEAR AND NONLINEAR IMPAIRMENTS IN FIBER-OPTIC COMMUNICATION SYSTEMS BY LDPC-CODED TURBO EQUALIZATION

by

Lyubomir Lyubomirov Minkov

Copyright © Lyubomir Lyubomirov Minkov 2009

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2009
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Lyubomir L Minkov entitled Experimental Demonstration of Mitigation of Linear and Nonlinear Impairments in Fiber-Optic Communication Systems by LDPC-coded Turbo Equalization and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Dr. Ivan B. Djordjevic  
Date: June 5, 2009

Dr. Ali Bilgin  
Date: June 5, 2009

Dr. Ali Akoglu  
Date: June 5, 2009

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director: Dr. Ivan B. Djordjevic  
Date: June 5, 2009
STATEMENT BY THE AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

Signed: ___________________________

Lyubomir L Minkov
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Ivan B. Djordjevic for the trust and helping me to bring out the best of myself at all times. I learned a great deal from working with you and most importantly to always expand the boundaries of my knowledge, find new ways to overcome obstacles, and bring the things I have started to completion. This work wouldn’t have been possible without your advice, support, help and faith in me. Thank you!

I would also like to thank,

Dr. Eniko Enikov for putting me on a path, that has lead to deep transformation and taught me to accept myself and others as they are.

Angelina Angelova for showing me how to follow the flow of life without resistance.

Hussam G. Batshon, for being a great friend and helping me see the world through the eyes of others.

Philip Georgiev and Christinka Gerogieva, for being next to me in my highs and lows and never giving up on me.
DEDICATION

This thesis is dedicated to my parents, Lyubomir and Diana, my brother and sister, Ivan and Sylvia for all the love, support and faith they have always had for me. I will always love you and know you will always be with me no matter where I go or what I do.
# TABLE OF CONTENTS

**LIST OF FIGURES** ................................................................................................................................. 8

**LIST OF TABLES** ................................................................................................................................. 11

**ABSTRACT** ........................................................................................................................................ 12

**CHAPTER I - INTRODUCTION** ......................................................................................................... 14

**CHAPTER II - FUNDAMENTALS OF OPTICAL COMMUNICATION SYSTEMS AND LDPC CODING**. 17

2.1 Introduction ..................................................................................................................................... 17

2.2 Modulation formats ....................................................................................................................... 17

2.2.1. NRZ format ............................................................................................................................ 17

2.2.2. RZ format ............................................................................................................................. 18

2.3 Linear impairments ....................................................................................................................... 19

2.4 Nonlinear impairments .................................................................................................................. 26

2.5 Coherent detection ....................................................................................................................... 30

2.6 LDPC codes ................................................................................................................................... 32

**CHAPTER III - CAPACITY CALCULATION** ....................................................................................... 45

3.1 Introduction ..................................................................................................................................... 45

3.2 Capacity calculation ....................................................................................................................... 47

3.3 Simulation results ........................................................................................................................... 49

3.4 Experimental results ..................................................................................................................... 55

3.5 Conclusion ...................................................................................................................................... 60

**CHAPTER IV - LDPC CODES FOR HIGH SPEED OPTICAL COMMUNICATION SYSTEMS** ............ 62

4.1 Introduction ..................................................................................................................................... 62
LIST OF FIGURES

Fig. 2.1 (a) NRZ TRANSMITTER CONFIGURATION (b) EXAMPLE FOR AN NRZ SIGNAL...... 18
Fig. 2.2 (a) RZ TRANSMITTER CONFIGURATION (b) EXAMPLE FOR A RZ SIGNAL.......... 19
Fig. 2.3 CHROMATIC DISPERSION EXAMPLE. ............................................................... 22
Fig. 2.4 FIBER BIREFRINGENCE. .................................................................................. 24
Fig. 2.5 DIFFERENTIAL GROUP DELAY BETWEEN THE AXES OF POLARIZATION IN FIBER... 25
Fig. 2.6 SECTIONS IN FIBER CONSIDERED FOR THE FOURIER SPLIT-STEP ALGORITHM...... 28
Fig. 2.7 (a) COHERENT DETECTION RECEIVER BLOCK DIAGRAM. (b) OPTICAL HYBRID..... 31
Fig. 2.8 BALANCED RECEIVER FOR HETERODYNE COHERENT DETECTION. ................. 32
Fig. 2.9 PARITY CHECK MATRIX AND THE CORRESPONDING BIPARTITE GRAPH. .......... 37
Fig. 2.10 (a) GIRTH FOR THE BIPARTITE GRAPH SHOWN IN Fig. 9 (b) EXAMPLES OR TYPICAL CONFIGURATIONS OF PARITY CHECK MATRICES RESULTING IN CODE GIRTHS OF G = 4
AND G = 6. .................................................................................................................. 38
Fig. 2.11 UPDATE RULES FOR THE ITERATIVE LDPC DECODING ALGORITHM. (a) UPDATES FOR THE CHECK NODES. (b) UPDATES FOR THE VARIABLE NODES. ...................... 38

Fig. 3.1 I.I.D. CHANNEL CAPACITY SIMULATIONS FOR DGD=100ps AND 10Gb/s RZ
SYSTEM......................................................................................................................... 50
Fig. 3.2 I.I.D. CHANNEL CAPACITY SIMULATION FOR 10Gb/s NRZ SYSTEM AND 700 OF
SMF28. .................................................................................................................................. 51
Fig. 3.3 I.I.D. CHANNEL CAPACITY SIMULATION FOR 10Gb/s NRZ SYSTEM AND 700 OF
SMF28 AND 50ps OF DGD. ............................................................................................. 51
Fig. 3.4 I.I.D. CHANNEL CAPACITY AS A FUNCTION OF THE NUMBER OF LINKS/SPANS IN A 40Gb/s SYSTEM THAT CONSISTS OF MULTIPLE LINKS. THE DISPERSION MAP IS SHOWN IN Fig. 3.5.......................................................................................................................... 52
Fig. 3.5 DISPERSION MAP FOR DOMINANT INTRA-CHANNEL FOUR WAVE MIXING......... 53
Fig. 3.6 I.I.D. CHANNEL CAPACITY LOSS AS A RESULT OF LLR QUANTIZATION AS A FUNCTION OF THE NUMBER OF SPANS FOR THE DISPERSION MAP OF Fig. 16 AND RZ TRANSMISSION AT 40Gb/s. .................................................................................................................. 54
Fig. 3.7 CONDITIONAL PDFS FOR TWO STATES OF THE BCJR TRELLIS IN THE PRESENCE
OF DGD.................................................................................................................................... 55
Fig. 3.8 EXPERIMENTAL SETUP. ..................................................................................... 56
LIST OF FIGURES - CONTINUED

FIG. 3. 9 MEMORY EFFECTS ON I.I.D. CHANNEL CAPACITY: (a) BER PLOTS AT MEMORY $2m+1=5$, (b) BER PLOTS AT MEMORY $2m+1=1$. ................................................................. 57

FIG. 3. 10 QUANTIZATION EFFECTS INFLUENCE ON I.I.D. CHANNEL CAPACITY (C): (a) LLR QUANTIZATION, (b) A/D CONVERSION ................................................................. 59

FIG. 4. 1 EXPERIMENTAL SETUP. ........................................................................................................ 76

FIG. 4. 2 PERFORMANCE EVALUATION OF LDPC CODES WITHOUT IMPAIRMENTS FOR DIFFERENT VALUES OF OSNR .................................................................................. 78

FIG. 5. 1 EXAMPLE OF CHANNEL DESCRIPTION WITH TRELlis THE TIME PERIODS $J$ AND $J-1.83$
FIG. 5. 2 FORWARD AND BACKWARD RECURSIONS FOR THE BCJR EQUALIZER ................. 85
FIG. 5. 3 BLOCK DIAGRAM OF PROPOSED TURBO EQUALIZATION SCHEME .................. 87
FIG. 5. 4 (A) DISPERSION MAP UNDER STUDY OF CHROMATIC DISPERSION COMPENSATION VIA LDPC-CODED TURBO EQUALIZATION (B) PERFORMANCE OF THE TURBO EQUALIZER FOR CHROMATIC DISPERSION COMPENSATION FOR LDPC(16935,13550) CODE, MEMORY $2m+1=7$ AND 700KM OF SMF28. ...................................................... 90
FIG. 5. 5 DISPERSION MAP FOR DOMINANT INTRA-CHANNEL FOUR WAVE MIXING ............. 91
FIG. 5. 6 BER PERFORMANCE OF TURBO EQUALIZATION FOR DIFFERENT CODING SCHEMES WITH THE DISPERSION MAP FROM FIG. 28 PERFORMANCE OF LARGE GIRTH CODES WITH NO EQUALIZER VS BCH(128,113)xBCH(256,239). ......................................................... 92
FIG. 5. 7 BER PERFORMANCE FOR DIFFERENT CODING SCHEMES FOR THE DISPERSION MAP ON FIG. 5. 5 LARGE GIRTH LDPC CODES VS TURBO-PRODUCT CODES AND LDPC WITH EQUALIZATION ................................................................................................................. 93
FIG. 5. 8 EXPERIMENTAL SETUP. ......................................................................................................... 94
FIG. 5. 9 PERFORMANCE EVALUATION OF LDPC CODES IN THE PRESENCE OF CONSTANT AMOUNT OF DGD (125PS) .................................................................................................................. 96
FIG. 5. 10 PERFORMANCE EVALUATION OF LDPC(11936,10819) CODE OF RATE $R = 0.906$, FOR DIFFERENT VALUES OF DGD ................................................................. 97
LIST OF FIGURES - CONTINUED

FIG. 6.1 Polarization multiplexed multilevel LDPC-coded modulation: (A) transmitter architecture, and (B) receiver architecture. CWL: continuous wave laser, PBS: polarization beam splitter, PBC: polarization beam combiner. ................................................................................................................. 104

FIG. 6.2 (A) 4-level trellis example for channel description, at time instant J. (B) constellation diagram for the used modulation format. .......................... 107

FIG. 6.3 (A) Forward recursion calculation for 4-level BCJR equalizer (B) Backward recursion calculation for 4-level BCJR equalizer. ...................... 109

FIG. 6.4 Experimental setup for polarization multiplexed BPSK. CW laser: continuous wave laser, PM: phase modulator, ASE: amplified spontaneous emission noise source, 3dB: 3dB coupler.......................................................... 113

FIG. 6.5 BER performance of proposed polarization multiplexed LDPC-coded modulation scheme. .............................................................................................. 115

FIG. 6.6 BER performance of separate polarization channels utilizing LDPC-coded turbo equalization: (A) DGD=0 ps, and (B) DGD=100 ps. .............. 116
LIST OF TABLES

TABLE 1 INVESTIGATED LDPC codes ............................................................... 77
TABLE 2 LDPC codes tested with turbo equalizer ......................................... 95
ABSTRACT

The ever-increasing demands for transmission capacity are the cause for the quick evolution of optical communication systems. Channel transmission at 100 Gb/s is already being considered by network operators. The major transmission impairments at these high rates are intra-channel and inter-channel nonlinearities, nonlinear phase noise, and polarization mode dispersion. By implementing LDPC-coded modulation schemes with soft decoding and Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm for equalization we have demonstrated significant improvements in system performance experiencing several impairments simultaneously. The new turbo-equalization scheme is used as a mean to simultaneously mitigate both linear and nonlinear impairments. This approach is general and applicable to both direct and coherent detection.

We provide comprehensive study of LDPC codes suitable for implementation in high-speed optical transmission systems. We determine channel capacity based on the forward step of the BCJR algorithm and show that by using LDPC codes we can closely approach the maximum transmission capacity that is possible. We propose the multilevel maximum a posteriori probability (MAP) turbo equalization scheme based on multilevel BCJR algorithm and an LDPC decoder, which considers independent symbols transmitted over both polarizations as two dimensional super-symbols. The use of multilevel modulation schemes provide higher spectral efficiency, while all related signal processing is performed at lower symbol rates, where dealing with PMD compensation
and fiber nonlinearities mitigation is more manageable. We show significant improvement in system performance over a system employing an equalizer that considers symbols transmitted in different polarizations as independent.
CHAPTER I - INTRODUCTION

This dissertation is dedicated to improving the performance of high-speed optical communication systems. The ever-increasing demands for transmission capacity require the development of novel concepts and methods for system improvement. The network operators already consider 100 Gb/s per wavelength transmission. At 100 Gb/s, optical fiber communications with conventional technologies face the technical challenges of strong signal degradations caused by transmission impairments, such as intra- and inter-channel nonlinearities, the nonlinear phase noise, and polarization-mode dispersion (PMD) [1]-[3]. In order to mitigate the signal distortions at ultra-high bit rates, some new technologies have been proposed and deployed in optical systems, and they represent a distinctive new trend in optical fiber communications. These new technologies include digital signal processing (DSP)-aided optical channel equalization, digital coherent receiving, multilevel modulations and optical polarization multiplexing (or optical multiple input multiple output technologies) [4]-[13].

Forward error correction (FEC) coding has become an efficient approach for performance enhancement in optical communication systems. The codes used for FEC in the past are primarily based on hard decision codes. Several soft-decision coding schemes with iterative decoding have been investigated recently, such as turbo product codes [14]-[17]. The most common equalization scheme used in the past is the maximum likelihood sequence detection (estimation) MLSD(E) based on Viterbi algorithm [18]-[20], which is
called here Viterbi equalizer. Unfortunately, the Viterbi equalizer provides hard decisions, and as such is not compatible with soft-decoding FEC schemes. We propose the use of structured low-density parity-check (LDPC) codes specifically suited for high-speed optical communication systems. They are soft-decision codes with iterative decoding, and for reasonable complexity they can approach the Shannon capacity limit within 1 dB. We also propose a novel LDPC-coded turbo equalization scheme to take the full advantage of the soft-decision information used in LDPC decoding. To improve the tolerance to fiber nonlinearities and PMD we perform the iteration of extrinsic information between BCJR equalizer and LDPC decoder. The proposed turbo equalization scheme is further generalized and proposed for use with multilevel modulation formats, which increases spectral efficiency and enables the use of more advanced modulation techniques. In addition, we propose an accurate method for channel capacity estimation that considers the interaction of different and nonlinear channel impairments. We show that with proposed turbo equalization scheme we can closely approach the channel capacity.

The rest of this work is organized as follows. In Chapter 2 we describe basic principles of optical transmission systems and LDPC coding. Short descriptions of linear and nonlinear signal impairments are provided. The non-linear Schrödinger equation and its numerical solution are reviewed. The sum-product-with-correction-term algorithm for decoding of LDPC codes is described. Chapter 3 describes the proposed method for channel capacity calculation in detail. It also provides simulation results and experimental results demonstrating the concept. Chapter 4 discusses the construction methods for
several classes of structured LDPC codes. LDPC codes suitable for implementation in high-speed optical transmission systems are proposed and experimental verification of their performance is provided. Chapter 5 introduces the concept of turbo equalization based on binary maximum a posteriori probability (MAP) detector implementing using Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm (called here BCJR equalizer) and LDPC decoder. The binary BCJR equalizer is fully described. Simulation results for several high-speed optical transmission systems in the presence of nonlinearities are provided to verify the concept of turbo equalization. In addition, we provide experimental verification results for systems in which the PMD is a predominant channel impairment. Chapter 6 generalizes the concept of the binary turbo equalization scheme proposed in Chapter 5 and applies it to multilevel modulation formats. The principle of operation of the multilevel BCJR equalizer is discussed in full detail. Experimental demonstration for a polarization multiplexed multilevel modulation scheme in the presence of PMD is provided. In final chapter, we provide some important concluding remarks and an overview of future work.
CHAPTER II - FUNDAMENTALS OF OPTICAL COMMUNICATION SYSTEMS AND LDPC CODING

2.1 Introduction

In this introductory chapter we discuss some modulation formats, introduction to LDPC codes and decoding, signal impairments for high-speed optical communication systems, coherent detection and optical pulse propagation in fiber.

2.2 Modulation formats

Modulation formats for optical communication systems fall into four general categories. amplitude shift keying (ASK), phase shift keying (PSK), frequency shift keying (FSK), and polarization shift keying (PolSK). The simplest formats used are non-return to zero (NRZ) ASK, and return to zero (RZ) ASK. We will give a brief description of these two formats next.

2.2.1. NRZ format

This format has the lowest bandwidth requirements. The modulator has to be of bandwidth 70% the bandwidth of the signal rate. This format has been widely used for years and is currently used primarily for comparison purposes with other, more advanced modulation formats.
Fig. 2.1 (a) shows the transmitter configuration for the NRZ format. It consists of a continuous wave laser and external modulator. In most applications Mach-Zehnder Modulator (MZM) is used [21]. It modulates the intensity of the laser at maximum or minimum for every pulse period. Maximum intensity pulses represent binary 1s and minimum intensity pulses represent binary 0s. The signal pattern generated in this fashion is called on-off keying (OOK) and an example is shown in Fig. 2.1 (b).

![NRZ transmitter configuration](image1)

**Fig. 2.1** (a) NRZ transmitter configuration (b) Example for an NRZ signal

### 2.2.2. RZ format

The return to zero format is similar to the NRZ in terms of signal generation, but it requires more bandwidth because for every bit period of time the signal returns to zero regardless if it was a 0 or a 1. The ratio of the pulse width at $\sqrt{2}/2$ of the maximum and the bit time period $T$ is the duty cycle for RZ modulation. As a result less power required for this format. Typically duty cycle of 33% is used. The bandwidth required for the RZ format is equal to the signal data rate. An advantage of the RZ scheme is that it is self-clocking and it is easier to synchronize.
Fig. 2.2 (a) shows the transmitter configuration for an RZ ASK signal. Fig. 2.2 (b) gives an example of such signal. To create an RZ signal two modulators are needed. The first one operates at twice the rate of the transmission rate. It creates a pulse train which is then modulated by the second modulator at the transmission rate. Synchronization between the pulse train and the second modulator is required for proper operation.

![Diagram of RZ transmitter configuration](image)

**Fig. 2.2** (a) RZ transmitter configuration (b) Example for a RZ signal.

### 2.3 Linear impairments

Optical transmission systems use fiber as a waveguide. For the purpose of explanation of the principle of operation of such systems, we can assume that the fiber is a linearly behaving media. These assumptions include that refractive index is not a function of power, the superposition principle of signals is applicable, wavelengths are not experiencing perturbations during transmission, and that the different signals in the same fiber do not interact [24]. If these assumptions are met the only impairments in an optical transmission system are the linear ones.

A pulse propagating in a fiber consists of multiple monochromatic waves propagating simultaneously. Generally each axial component of the monochromatic electromagnetic wave can be by expressed by its complex electric field function [22]
where the propagation is over the $z$ axis of the coordinate system attached to the fiber.

The parameter $\beta$ is the propagation constant of the fiber and is an essential parameter for the effects on pulse propagation. The Taylor expansion of $\beta$ is essential to understanding pulse propagation [23].

\[
\beta(\omega) = \beta(\omega_0) + \frac{d\beta}{d\omega}(\omega - \omega_0)_{|\omega=\omega_0} + \frac{1}{2} \frac{d^2\beta}{d\omega^2}(\omega - \omega_0)^2_{|\omega=\omega_0} + \frac{1}{6} \frac{d^3\beta}{d\omega^3}(\omega - \omega_0)^3_{|\omega=\omega_0} + \ldots
\]

The first three components of the Taylor expansion have the following meaning

\[
\beta_1 = \frac{d\beta}{d\omega} = \frac{1}{v_g}, \text{ where } v_g \text{ is the group velocity of the pulse}
\]

\[
\beta_2 = \frac{d^2\beta}{d\omega^2} \text{ is the group velocity dispersion (GVD) parameter}
\]

\[
\beta_3 = \frac{d^3\beta}{d\omega^3} \text{ is the second order group velocity dispersion parameter}
\]

The group velocity of the optical pulse represents the speed at which the energy of the pulse propagates through the medium. The other two components are related to the dispersion of the signal in the fiber.

After substituting the expansion of the propagation parameter in the propagation equation, the basic form of the propagation equation of an optical pulse in single mode fiber is achieved [23]
To fully understand the effects of the linear impairments on the propagation a brief description of them is given next. They include fiber attenuation, frequency chirp, chromatic dispersion, first order polarization mode dispersion and impairments caused by imperfect components like insertion loss.

The fiber attenuation is given by the power attenuation coefficient $\alpha$. It occurs because of material imperfections. Rayleigh and Mi scattering occur in the fiber during transmission as well as intrinsic absorption due to resonances in the fiber close to transmission wavelength and extrinsic absorption due to impurities in the material. The fiber attenuation coefficient provides the relationship between the output power and the input power after a transmission distance $L$ kilometers $P_{\text{out}} = P_{\text{in}} e^{-\alpha L}$. The attenuation constant is often used in dB/km $\alpha[\text{dB/km}] \approx 4.343 \cdot \alpha[1/\text{km}]$.

The frequency chirp occurs at the transmitter end of an intensity modulated optical system. When pulses are generated, intensity modulation causes phase modulation due to the carrier-induced change in the refractive index. This change is inherent because of the laser linewidth. Optical pulses with a time-dependent phase shift are called chirped pulses. The chirp results in broadening of the optical spectrum of the transmitted signal. Frequency chirp can also occur due to self-phase modulation of the signal.

Chromatic dispersion is the most common problem experienced during transmission. A perfect laser source would contain only a single frequency, however real
devices have laser linewidth and upon modulation pulse contains multiple spectral components. Different spectral components of the optical pulse experience different refractive index during propagation. At launch they are aligned but with time they propagate with different speeds causing some of them to arrive ahead of others. In time domain this is observed as pulse broadening. This is shown in Fig. 2.3.

This effect can be described by the pulse broadening $\Delta \tau_g = (d\tau_g / d\lambda)\Delta \lambda$, where $\tau_g = L / v_g = L\beta_l$ is the time delay that a spectral component experiences after travelling for distance $L$. It is also known as group delay. $\lambda$ is the wavelength and $c$ is the speed of light. $\Delta \lambda$ is the range of wavelengths emitted by the optical source.
The pulse broadening can be represented as

$$\Delta \tau_g = L \frac{d \beta_1}{d \lambda} \Delta \lambda = \frac{2 \pi c}{\lambda^2} \beta_2 L \Delta \lambda = DL \Delta \lambda$$

(2.3)

$D$ is the chromatic dispersion parameter. It can be represented by two components $D = D_M + D_W$.

$D_W = \frac{2}{\lambda} \beta_1$ is the waveguide dispersion. It occurs because the propagation constant is a function to wavelength $\beta = \beta(\lambda)$ and also depends on the geometry of the fiber. $D_W$ is related to the slope of the group refraction index which is defined as $n_g = n - \lambda \frac{dn}{d\lambda}$.

Typical value of dispersion coefficient for SMF28 is $D=15-18 \text{ps/(km-nm)}$.

$D_M = -\frac{2 \pi c}{\lambda^2} \beta_2$ is the material dispersion. It is caused by wavelength dependence of the refractive index in the fiber $n = n(\lambda)$. This results in wavelength dependence of the group delay.

Higher order dispersion is determined by the derivative of the dispersion in terms of wavelength. It is called dispersion slope parameter, also known as differential dispersion parameter, or second-order dispersion parameter and is related to the first and second order GVDs.
Another form of dispersion in optical fiber is the polarization mode dispersion (PMD). It is caused by birefringence or the fact that in the optical fiber the refractive index is not constant for the different polarizations and it varies along the fiber due to fiber imperfections and stress in the fiber caused from external or internal factors, for example bending and material impurities.

The PMD is determined by the degree of birefringence of the fiber $B_m = |\bar{n}_x - \bar{n}_y| = \Delta \bar{n}$, where $\bar{n}_x$ and $\bar{n}_y$ are the refractive indexes along the optical axes of the fiber also known as mode indexes for the orthogonally polarized fiber modes. The propagation constant is determined by the refractive index $\Delta \beta = |\beta_x - \beta_y| = (\omega / c) \Delta \bar{n}$, which defines the beat length of the birefringence $L_B = 2\pi / \Delta \beta = \lambda / B_m$. Fig. 2.4 illustrates the birefringence in the optical fiber. The state of polarization changes along the fiber in a periodic fashion.

\[
S = \frac{dD}{d\lambda} = \frac{4\pi c}{\lambda^3} \beta_2 + \left(\frac{2\pi c}{\lambda^2}\right)^2 \beta_3. \tag{2.4}
\]

---

**Fig. 2.4** Fiber birefringence.

The differential group delay (DGD) is defined as
It represents the delay between the two axes of polarization after propagation of some distance $L$. Fig. 2.5 illustrates the DGD for a propagating pulse in a fiber.

\[
\Delta \tau = \left| \frac{L}{v_{gx}} - \frac{L}{v_{gy}} \right| = L \left| \frac{d\omega}{\beta_x} - \frac{d\omega}{\beta_y} \right|
\] (2.5)

**Fig. 2.5** Differential group delay between the axes of polarization in fiber.

DGD is a random variable with Maxwellian distribution with mean-square value of [24]

\[
\sigma^2(z) = \left\langle (\Delta \tau)^2 \right\rangle = \left( \Delta \beta \sqrt{2l_c L} \right)^2,
\] (2.6)

where $l_c$ is the correlation length. This is the length after which the two polarizations can be considered decoupled. For long fiber $L >> l_c$ the mean-square value of DGD can be written as [24]

\[
\left\langle (\Delta \tau)^2 \right\rangle^{1/2} = \Delta \beta \sqrt{2l_c L} = D_p \sqrt{L}
\] (2.7)

$D_p$ is the PMD parameter. Typical values for $D_p$ are in the range 0.01-10ps/(km)$^{1/2}$. This is the first order PMD. It interacts with the chromatic dispersion and can be considered as a component of chromatic dispersion. The first order PMD is characterized by a highly
asymmetric Maxwellian probability density function (PDF). It can be used to determine
the mean value of $D_p$.

The second order PMD is denoted by $D_{p2}$ and can be determined from

$$D_{p2} = \left[ \frac{1}{2} \frac{\partial D_p}{\partial \omega} + \left( \frac{D_p}{2} \frac{\partial |s|}{\partial \omega} \right)^2 \right]^{1/2},$$

(2.8)

where $s$ is the Stokes vector for the polarization. The mean value of $D_{p2}$ can be related to the
mean value of $D_p$. This term is related to the rotation of the principle axes of polarization
in the fiber and is highly nonlinear thus cannot be treated as a component of chromatic
dispersion.

2.4 Nonlinear impairments

The assumptions for linear regime of operation in fibers are not true and eventually
must be accounted for in order to improve system performance. Modern systems have to
meet high bandwidth requirements and dense wavelength division multiplexing is being
pushed to its limits. The presence of multiple channels with marginal spacing and small
tolerances are the cause of signal interaction within the fiber and optical power exceeding
the minimum required for neglecting nonlinearities. The two major types of nonlinearities
are the Kerr nonlinearities related to the nonlinear behavior of the refractive index, and
the nonlinear optical scattering. To reflect these nonlinearities the linear propagation
equation (2) can be rewritten in more general form to include nonlinear terms [23]

$$\frac{\partial A(z,t)}{\partial z} = -\frac{\alpha}{2} A(z,t) - \beta_1 \frac{\partial A(z,t)}{\partial t} - j \frac{\beta_2}{2} \frac{\partial^2 A(z,t)}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 A(z,t)}{\partial t^3} + j \gamma |A(z,t)|^2 A(z,t),$$

(2.9)
After transformation to a new coordinate system the equation is reduced by one term and thus simplified. For \( T = t - \beta z \) the equation becomes

\[
\frac{\partial A(z,T)}{\partial z} = -\frac{\alpha}{2} A(z,T) - j\frac{\beta_2}{2} \frac{\partial^2 A(z,t)}{\partial T^2} + \frac{\beta_3}{6} \frac{\partial^3 A(z,t)}{\partial T^3} + j\gamma |A(z,T)|^2 A(z,T) \quad (2.10)
\]

and is known as nonlinear Schrödinger equation.

\[
\gamma = \frac{2\pi n_2}{\lambda A_{\text{eff}}} \quad \text{is the nonlinear coefficient.}
\]

\( n_2 \) is the nonlinear Kerr coefficient.

\( A_{\text{eff}} \) is the effective cross-section area of the optical fiber core.

The most common method for solving the nonlinear Schrödinger equation is the Fourier split-step algorithm briefly described next. First a linear operator is defined in time domain and in frequency domain after Fourier transform.

\[
D = -\frac{\alpha}{2} - j\frac{\beta_2}{2} \frac{\partial^2}{\partial T^2} + \frac{\beta_3}{6} \frac{\partial^3}{\partial T^3} \quad \text{and} \quad \tilde{D} = -\frac{\alpha}{2} - j\frac{\beta_2}{2} \omega^2 + \frac{\beta_3}{6} \omega^3 \quad (2.11)
\]

The nonlinear part of the nonlinear Schrödinger equation is denoted as a nonlinear operator \( N = j\gamma |A(z,T)|^2 \). Both the linear and the nonlinear parts of the equation have analytical solutions, but no general analytical solution exists for them simultaneously. The equation can be numerically solved by iterative steps by using solving for regular intervals over the fiber. Assuming that the wave is known at the beginning of the fiber \( z_0 \), the fiber is divided in sections of sufficiently small length \( h \) to ensure that the two parts of the equation can be treated as independent of each other over that length. Then the linear
and non-linear parts of the equation can be solved for separately with negligible error. For the first half of every interval $h$ only the linear part of the equation is considered. Fast Fourier transform is performed and the linear operator is applied in frequency domain. The result is transformed back into time domain via inverse Fourier transform and now only the nonlinear part of the equation is considered. At the middle of the interval the inverse Fourier transform is used and the linear operator is applied in time domain. At this point the equation is again transformed in frequency domain and solved for in frequency domain for the remaining half interval $h/2$.

Fig. 2.6 shows the sections in the fiber considered for that solution.

![Diagram](image)

**Fig. 2.6** Sections in fiber considered for the Fourier split-step algorithm.

This procedure can be written in operator form as follows

$$A(z_0 + h, T) = e^{hD/2} [e^{hN} A(z_0, T)] e^{hD/2}.$$  

A brief description of the fiber nonlinearities is provided next. As mentioned the major types of fiber nonlinearities are the Kerr nonlinearities and the simulated scattering nonlinearities.
The Kerr nonlinearities include self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM). SPM occurs at high intensity of the optical signal. The refraction index of the fiber is dependent on the intensity of the optical pulse and different sections of the pulse in time domain experience different refractive indexes leading to chirping of the pulse. The same nonlinearity is the cause of the XPM. It is related to multichannel transmission at different wavelengths in the same fiber or composite signal propagation. The adjacent channels affect each other by introducing nonlinear phase shift for the periods of time where the signals overlap. The leading edge of the pulses experience a decrease in optical frequency (red shift) and the trailing edges of the pulse experience an increase in frequency (blue shift). The pulses become chirped, this causes them to overlap and intersymbol interference occurs.

FWM occurs in a propagation of composite signal as well. When a phase matching condition is met, three optical signals with different carrier frequencies interact and a fourth, previously non-existing, signal is generated. The power of the newly generated frequency depends on the powers of the three signals that generated it, the intensity of the Kerr nonlinear effect, and the satisfaction of the phase-matching condition. Significant degradation in WDM systems can occur because the newly generated frequencies can coincide with an existing channel in the composite signal and cause the appearance of ghost pulses and intersymbol interference.

The scattering effects that cause nonlinear signal distortion are the stimulated Raman scattering (SRS) and the stimulated Brillouin scattering (SBS) [24]. The SRS
occurs when the propagating signal interacts with the glass molecules of the fiber. This interaction causes a wavelength shift of the signal which can result in adjacent channels interacting with each other and resulting in power transfer from higher frequencies in the composite signal to the lower ones. The channel with shortest wavelength serves as a pump to several of the channels with longer wavelengths and undergoes depletion in power.

The SBS occurs when the propagating optical signal interacts with acoustic phonons [24]. The incident optical signal is reflected from the grating formed by the acoustic vibrations and a downshift in frequency occurs. Acoustic vibrations originate from thermal effects in the lattice of the glass caused by the power of the signal. The periodically changing material density generates acoustic waves. High power signal causes the effect to be stimulated.

### 2.5 Coherent detection

Optical communication systems with direct detection are most widely used. Direct detection only registers the intensity of the electrical field and doesn’t provide phase information about the signal. Many modulation formats and more elaborate setups are impossible without the presence of phase information about the signal. Coherent detection provides the phase information about the signal, increases the receiver sensitivity, better frequency selectivity, enables many formats that utilize constant amplitude and have the advantage of tunable optical receiver. There are two major reception schemes for coherent reception – homodyne and heterodyne. The detection can
be synchronous or asynchronous. The principle of operation of coherent receiver is shown in Fig. 2.7 (a) [22].

![Diagram of coherent detection receiver](image1)

**Fig. 2.7** (a) Coherent detection receiver block diagram. (b) Optical hybrid.

The basic idea of coherent detection is the presence of a second laser signal at the receiver. It is called local laser and it operates at the same wavelength as the signal carrier. The two signals are mixed at the receiver before detection. A 4 port device used for mixing of the signals and it is called optical hybrid. A simplified optical hybrid diagram is shown in Fig. 2.7 (b). The outputs of the optical hybrid are related to the inputs of the electrical fields as follows:

\[
E_{io} = (E_{ii} + E_{i2})\sqrt{1-k} \\
E_{2o} = (E_{ii} + E_{i2}e^{-j\phi})\sqrt{1-k},
\]  

where \(\phi\) is the phase correction introduced by the controlled voltage of the hybrid. The hybrid is characterized by the \(S\)-matrix of a 4 port device.

\[
S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{21} \end{bmatrix}, \quad S = \begin{bmatrix} \sqrt{1-k} & \sqrt{1-k} \\ \sqrt{k} & e^{-j\phi}\sqrt{k} \end{bmatrix}, \quad S_x = \begin{bmatrix} \sqrt{1/2} & \sqrt{1/2} \\ \sqrt{1/2} & -\sqrt{1/2} \end{bmatrix}
\]  

(2.13)
Different values of $k$ and $\phi$ result in different types of hybrids. A popular one is the $\pi$-hybrid. An example of coherent receiver with $\pi$-hybrid and a balanced detector is shown in Fig. 2.8 [25].

![Balanced receiver for heterodyne coherent detection.](image)

**Fig. 2. 8** Balanced receiver for heterodyne coherent detection.

The signals in the balanced receiver are

$$i_1(t) = R \left| E_1 \right|^2 = \frac{1}{2} R (P_S + P_{LO} + 2\sqrt{P_S P_{LO}} \cos(\theta_S)) + n_1(t)$$

$$i_2(t) = R \left| E_2 \right|^2 = \frac{1}{2} R (P_S + P_{LO} - 2\sqrt{P_S P_{LO}} \cos(\theta_S)) + n_2(t)$$

$$i(t) = i_1(t) - i_2(t) = 2\sqrt{P_S P_{LO}} \cos(\theta_S) + n(t)$$

(2.14)

where $n_1(t), n_2(t), n(t)$ are the noise signals for the two diodes and for the whole detector respectively. $\theta_S$ is the phase of the incoming signal.

### 2.6 LDPC codes

Low density parity check (LDPC) codes are linear block codes (LBC) with very sparse parity check matrices. They were initially introduced by R. Gallagher in [26]. He proposed codes based primarily on semi-random construction. The first graphical representation of LDPC codes was provided by R. Tanner [27]. He also generalized these
codes and made the first step towards systematic code design. Different code designs have been proposed over the years [28]-[34]. Two important features of the LDPC codes are that they can be iteratively decoded and the use soft information for decoding.

LDPC codes are linear block codes [35]. A linear block code operates on the output of an information source which is a binary information sequence. This sequence is segmented into message blocks of equal length $k$. Each block is denoted by $m$ and there are exactly $2^k$ distinct messages. The LBC serves as an encoder, which according to certain rules, transforms each input message $m$ into a binary $n$-tuple $c$ called a codeword, where $n > k$. The set of $2^k$ codewords corresponding to the $2^k$ possible messages is called a block code. Typical notation for LBCs is $(n,k)$. The LBC can be viewed as an $n$-dimensional space of the binary field. This subspace is uniquely mapped to all possible $2^k$ binary messages. Matrix multiplication represents the encoding process in a straightforward fashion. A matrix of dimensions $(n,k)$ will serve as an encoder provided that the elements of the matrix are in the set $\{0,1\}$. The elements of the output of the multiplications of any binary message with the so described encoder matrix remain in the binary set $\{0,1\}$. The common notation accepted for the encoder matrix is $G$ and it is called generator matrix. The rows of the matrix are the set of vectors $\{g_0, g_1, \ldots, g_{k-1}\}$. They are binary linearly independent vectors of length $n$ and form an orthogonal basis which spans the whole $n$-dimensional space of the binary field. A null space exists for the generator matrix. It consists of $r$ linearly independent binary vectors of length $n$ with the property that every vector from the null space is orthogonal to every vector in the
generator matrix. Here \( r = (n-k) \). The ordered set of vectors of the null space are denoted as \( \{ h_0, h_1, \ldots, h_{n-k-1} \} \). They form a second matrix denoted by \( H \) also known as a parity-check matrix as given by (15). The definition of the parity-check matrix implies that \( H^T G = 0 \)

\[
G = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{n-k-1} \end{bmatrix}, H = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{n-k-1} \end{bmatrix}
\]

(2.15)

A popular representation of \( G \) and \( H \) exist, which is shown in (16). It is called systematic form of the generator and parity-check matrices.

\[
G = \begin{bmatrix} P_{k \times (n-k)} | I_{k \times k} \end{bmatrix} = \begin{bmatrix} P_{00} & P_{01} & \cdots & P_{0,n-k-1} \\ P_{10} & P_{11} & \cdots & P_{1,n-k-1} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k-1,0} & P_{k-1,1} & \cdots & P_{k-1,n-k-1} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}
\]

(2.16)

\[
H = \begin{bmatrix} I_{(n-k) \times (n-k)} | P^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & P_{00} & P_{01} & \cdots & P_{0,k-1} \\ 0 & 0 & \cdots & 0 & P_{01} & P_{11} & \cdots & P_{1,k-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & P_{0,n-k-1} & P_{1,n-k-1} & \cdots & P_{k-1,n-k-1} \end{bmatrix}
\]

\( P \) is a submatrix that adds parity check information to the sequence of message bits, and \( I \) is the identity matrix. It is also called message part of the matrix because it repeats the original message in the resulting codeword sequence of \( n \) bits. As mentioned above the generation of a codeword \( c \) is simple matrix multiplication of the binary message vector \( m = [m_0, m_1, \ldots, m_{n-k-1}] \) with the matrix \( G \), \( mG = c \). The structure of the codeword is of the
form \( e = [c_0, c_1, ..., c_{n-1}] \). The last \( k \) bits in the codeword are identical to the original message. The first \( r=n-k \) bits in the codeword are the parity-check bits. When in systematic form the codeword can be written as \( e = [c_0, c_1, ..., c_r, m_0, m_1, ..., m_{k-1}] \). The first \( r \) bits are parity check bits and the remaining \( k \) bits are message bits. When generating a codeword, the parity-check bits can be represented as the set of parity-check equations:

\[
\begin{align*}
  c_0 &= m_0 p_{00} + m_1 p_{10} + m_2 p_{20} + \ldots + m_{k-1} p_{k-1,0} \\
  c_1 &= m_0 p_{01} + m_1 p_{11} + m_2 p_{21} + \ldots + m_{k-1} p_{k-1,1} \\
  &\vdots \\
  c_r &= m_r p_{0r} + m_1 p_{1r} + m_2 p_{2r} + \ldots + m_{k-1} p_{k-1,r}
\end{align*}
\]

(2.17)

They have to be satisfied for the codeword to be valid without error. Upon receiving a codeword, the parity-check matrix can be used for decoding. If the received codeword is valid, multiplication by \( H \) has to result in a zero vector because \( H \) is the null space of \( G \).

Let the binary vector \( e \) denote the result of such multiplication \( cH^T = e \). Note that this multiplication results in the set of parity-check equations (17). If \( e = 0 \) the received codeword is valid and no correction is needed because all parity-check equations are satisfied. If \( e \) is not the zero vector, syndrome decoding can be performed. That means that the valid codeword containing the smallest number of differences to the received codeword is chosen to substitute the received codeword. There are many disadvantages to such method of decoding. To overcome this iteratively decodable codes are proposed. Iterative decoding substantially increases the error-correction capabilities of the decoder. LDPC codes belong to this class of codes and are described next in more detail.
The structure of the parity-check matrix for an LDPC code has the following properties: They have a specific number of 1s per row and specific number of 1s per column denoted by $w_r$ and $w_c$ respectively. A regular LDPC code has the same number of 1s for all columns while an irregular LDPC code has different number of 1s in different columns. Usually $w_c$ is defined for a certain code and $w_r$ is determined by it as $w_r = w_c(n/r)$. To achieve high degree of sparsity the condition $w_c << r$ needs to be met. The code rate of an LDPC code is determined by $R=1 - w_c / w_r$ for full rank parity check matrix $H$.

The graphical representation of LDPC codes is very convenient for their understanding. The parity-check matrix can be represented in a graphical form. Every row of the parity-check matrix is associated with a parity-check equation, as explained earlier. Every column of the parity check is associated with a bit in the received sequence of $n$ bits. A graph structure is created containing two rows of $r$ and $n$ nodes. The two sets of nodes are denoted as $r$ check nodes and $n$ variable nodes. They are connected in relation to the parity check equation. The check nodes represent the rows and the variable nodes represent the columns. A connection between a check node and a variable node exists where a 1 is present in the parity check matrix. A graph structure connected in this fashion is denoted as a bipartite graph. The variable nodes correspond to the received $n$ samples of a codeword and the check nodes correspond to the parity check equations for the codeword. Fig. 2.9 shows the bipartite graph for the following parity-check matrix [27].
An important parameter for every LDPC code is the code girth. It can be defined using the bipartite graph as shown in Fig. 2.10 (a). A cycle is a closed loop defined by the edges in the graph. The girth of the code is the cycle of minimum number of edges in the graph. The girth can be detected in the parity check matrix as a specific configuration of bits. An example for two typical such configurations is shown in Fig. 2.10 (b). The girth of a code is typically denoted as $g$. 

**Fig. 2.9** Parity check matrix and the corresponding bipartite graph.

$$H = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1
\end{bmatrix}$$
Fig. 2. 10 (a) Girth for the bipartite graph shown in Fig. 9 (b) Examples or typical configurations of parity check matrices resulting in code girths of \( g = 4 \) and \( g = 6 \).

Fig. 2. 11 Update rules for the iterative LDPC decoding algorithm. (a) Updates for the check nodes. (b) Updates for the variable nodes.

The bipartite graph representation of LDPC codes is convenient for understanding the iterative decoding process. Let the vector of received samples is \( Y = \{ y_1, y_2, \ldots, y_n \} \).
Upon receiving the samples they are associated with the variable nodes of the graph $y_1 \to c_1, y_2 \to c_2, \ldots, y_n \to c_n$. The iterative process of updating the variable nodes consists of two steps per iteration [36]. At every step messages are passed from variable nodes to check nodes or from check nodes to variable nodes. This process is illustrated in Fig. 2.11 (a) and Fig. 2.11 (b) respectively. In the example of Fig. 2.11 (a) variable node $c_7$ sends a message to check node $f_4$, which is denoted by $q_{74}$. In the example of Fig. 2.11 (b) check node $f_1$ sends a message to variable node $c_4$ which is denoted by $r_{14}$. The message $q_{74}$ is determined by the product of the messages received by $c_7$ from all other check nodes connected to it: $q_{74} = r_{27}r_{37}$. The check node $f_4$ receives messages from all variable nodes connected to it and in the second step of the iteration sends out messages to all variable nodes connected to it. In the example in Fig. 4 (b) check node $f_1$ sends a message $r_{14} = (q_{11} + q_{21} + q_{31})f_1(c_1, c_2, c_3)$. The function of the check node can vary. Usually it is modulo 2 summation. The second step finishes with the update of the variable node which acquires the new value according to a decision rule. Typically it is a majority rule, which means that the variable node acquires the bit value indicated by the majority of messages that it has received at the end of the current iteration. In case no initial values exist for the messages, the first iteration starts with the variable nodes sending the value of the received sample.

In the general case the message passing rules can be written as [36]
\[ q_{ij} = \prod_{j' \in Q \setminus j} r_{j' i} \quad \text{and} \quad r_{ji} = \sum_{-\{i\}} f(I) \prod_{i' \in R \setminus i} q_{i' j} \quad (2.18) \]

\( R_j \setminus i \) is the set of all messages received from variable node \( c_j \) excluding the message from check node \( f_i \) to \( c_j \). \( Q \setminus j \) is the set of all messages received from check node \( f_i \) excluding the message from variable node \( c_j \) to \( f_i \). \( I \) indicates the set of arguments of the check node function \( n(f) \). The message passed to the variable check node \( j \) is a summation over all check nodes connected to that variable node excluding the check node \( j \). A message passed from a variable node \( c_j \) to a check node \( f_i \) contains information about the conditional probabilities \( P(c_j = 0 \mid y_j) \) and \( P(c_j = 1 \mid y_j) \). Because a single message is passed at every iteration the information for both probabilities is sent as a ratio. This ratio is defined as likelihood ratio. It indicates the more likely of the two events \( c_j \) being 0 or 1 provided that \( y_j \) has been received.

In soft decoding algorithms every received bit is being processed along with the likelihood reliability for that bit. Typically the reliabilities are calculated in the log domain and are called log likelihood reliabilities (LLRs). Equation (19) shows the definition of the LLR for a bit \( v_i, i = 1, 2, \ldots, n \).

\[ L(v_i) = \log \left( \frac{P(v_i = 0 \mid y_i)}{P(v_i = 1 \mid y_i)} \right) \quad (2.19) \]

If \( L(v_i) > 0 \) the bit \( v_i \) is most probably 0, while if \( L(v_i) < 0 \) the bit \( v_i \) is most probably 1. Higher value of the LLR indicates higher certainty about the decision for the bit.
A block diagram for the decoding of the LDPC codes is presented next. It is based on the sum-product with correction term algorithm [2]. The log domain operation has the advantages of using summations instead of multiplications which leads to greater numerical stability. The notation from [28] is used here.

\[ V_j = \{ \text{variable nodes } c, \text{ connected to check node } f_j \} \]

\[ V_{ji} = \{ \text{variable nodes } c, \text{ connected to check node } f_j \} \backslash \{ \text{variable node } c_i \} \]

\[ C_i = \{ \text{check nodes } f, \text{ connected to variable node } c_i \} \]

\[ C_{ij} = \{ \text{check nodes } f, \text{ connected to variable node } c_i \} \backslash \{ \text{check node } f_j \} \]

\[ P_i = P(c_i = 1|y_i) \]

\[ M_v(\sim i) = \{ \text{all messages from variable nodes except variable node } c_i \} \]

\[ M_c(\sim j) = \{ \text{all messages from check nodes except check node } f_j \} \]

\[ S_i = \text{the event that the check equations for check node } f_j \text{ are satisfied} \]

\[ q_{ij}(b) = P(v_i = b|S_i, y_i, M_c(\sim j)) \]

\[ r_{ji}(b) = P(\text{parity check equation } f_j \text{ is satisfied}| v_i = b, M_v(\sim i)) \]

The probabilities for the log domain will be denoted as follows:

\[ L(r_{ji}) = \log \left( \frac{r_{ji}(0)}{r_{ji}(1)} \right) \]
\[ L(q_{ij}) = \log \left( \frac{q_{ij}(0)}{q_{ij}(1)} \right) \]

The algorithm follows the steps:

1) **Initialization** – set \( L(q_{ij}) = L(v_i) \), where

\[ L(v_i) = 2 \frac{v_i}{\sigma^2}, \text{ for binary input additive white gaussian noise channel} \]

where \( \sigma^2 \) is the variance of the Gaussian distribution for the AWGN

\[ L(v_i) = \log \left( \frac{\sigma_i}{\sigma_0} \right) - \frac{(v_i - \mu_0)^2}{\sigma_0^2} + \frac{(v_i - \mu_1)^2}{\sigma_1^2}, \text{ for binary asymmetric AWGN channel} \]

where \( \sigma_0, \sigma_1, \mu_0, \mu_1 \) are the mean values and the variances for the optical Gaussian pulse that represents the bit \( j \); the indexes 0 and 1 indicate the transmitted bit.

2) **Calculate** \( L(r_{ji}) \)

\[ L(r_{ji}) = \left( \prod_{i \in R_j \setminus i} \alpha_{i,j} \right) \phi \left[ \sum_{i \in R_j \setminus i} \phi(\beta_{i,j}) \right] \]

\( \alpha_{ij} = \text{sign}(L(q_{ij})) \), \( \beta_{ij} = \left| L(q_{ij}) \right| \), and \( \phi(x) = -\log \tanh(x/2) = \log \left( \frac{e^x + 1}{e^x - 1} \right) \). Then

\[ L(r_{ji}) = L \left( \sum_{i \in R_j \setminus i} b_i' \right) = L(...) \oplus b_k \oplus b_l \oplus b_m \oplus b_n \ldots = \ldots + L_k + L_l + L_m + L_n + \ldots, \]
where $\oplus$ is modulo-2 addition, and $+$ is a function defined as follows

$$L_a \oplus L_b = \text{sign}(L_a) \text{sign}(L_b) \min(|L_a|, |L_b|) + \log \frac{1 + e^{-|L_a - L_b|}}{1 + e^{-|L_b - L_a|}}$$

The term $\log \left( (1 + e^{-|L_a + L_b|}) / (1 + e^{-|L_a - L_b|}) \right)$ is called correction factor and is usually approximated as

$$\tilde{s}(x, y) = \begin{cases} c, |x + y| < 2 \text{ and } |x - y| > 2|x + y| \\ -c, |x + y| > 2 \text{ and } |x + y| > 2|x - y| \\ 0, \text{ otherwise} \end{cases}$$

where $c$ is a correction factor constant.

3) Calculate $L(q_{ij})$

$$L(q_{ij}) = L(v_i) + \sum_{j' \in C_i \setminus j} L(r_{ji})$$

4) Calculate $L(Q_i) = \log Q_i(0) - \log Q_i(1)$

$$L(Q_i) = L(v_i) + \sum_{j \in C_i} L(r_{ji})$$

5) Decision step:

For every bit of a received codeword $c$ calculate

$$\hat{c}_i = \begin{cases} 1, L(Q_i) < 0 \\ 0, L(Q_i) > 0 \end{cases}$$
Conditions for termination of iterations are:

1) $\hat{c} H^T = 0$

2) Number of maximum iterations has been reached.
CHAPTER III - CAPACITY CALCULATION

3.1 Introduction

The intensive research of optical transmission systems in recent years has lead to increase in spectral efficiency through the use of coherent detection, multilevel coding and dense wavelength division multiplexing (DWDM). The need for bandwidth has resulted in higher data rates and nonlinear effects in fiber transmission systems have become dominant. Accurate and precise determination of the Shannon limit capacity for different channels is of utmost importance for the future of optical communications. It will allow for maximizing of the channel throughput and provide a sound estimate of the performance of different systems and popular modulation formats.

Multiple attempts have been made to estimate the channel capacity in DWDM systems and in optical transmission systems in general [37]-[41]. From the approaches used in the past one of the most common approach is considering the amplified spontaneous emission (ASE) from optical amplifiers as dominant source of noise. The effect of nonlinearities is usually accounted for but usually they are considered either as perturbation of a linear case or as multiplicative noise. Both of these approaches are inaccurate. The reliability of such results is not sufficient for future work because current and future systems operate at 40Gb/s and higher. In this regime the nonlinearities are the dominant source of transmission impairments and cannot be considered negligible.
Besides the strength of the considered nonlinearities the type is also important. At lower data rates the inter-channel nonlinearities are the dominant type while at 40Gb/s and above the intra-channel nonlinearities become dominant. The most typical intra-channel nonlinearities are four-wave mixing (FWM), self-phase modulation (SPM) and cross-phase modulation (XPM). Accurate channel capacity estimation is required for the design and implementation of systems operating at high speeds. It is desirable that these estimation methods are suitable for experimental validation for the purpose of testing the installed systems. For that purpose realistic simulations that take into account the combined effects of amplified spontaneous noise (ASE), Kerr nonlinearities, chromatic and polarization mode dispersion, and filtering effects are preferred [42],[43].

In this chapter the simulation and experimental results for achievable information rates for high-speed long-haul optical transmission systems are reported. An independent identically distributed (i.i.d.) information source was used. The nature of the source indicates that the distribution of the 0s and 1s of the information source is even which allows for accurate calculation of the capacity. The so achieved capacity is called i.i.d. capacity and it is a lower bound on the channel capacity. This indicates the highest possible capacity that can be achieved by implementing the best known techniques for detection, equalization and coding up to date. The technique used for calculation of the i.i.d. channel capacity follows the approach described in [6]. It consists of two steps: (i) determination of conditional probability density functions (PDFs) from experimentally collected histograms, and (ii) capacity calculation based on the forward step of the BCJR algorithm and is described in detail later in the chapter.
Simulations were performed by using the nonlinear Schrodinger equation for the optical pulse propagation. The equation is solved by using the Fourier split step algorithm described in [23]. The i.i.d. capacity studies for RZ system at 10Gb/s in the presence of polarization mode dispersion and for NRZ system at 10Gb/s with chromatic dispersion are presented. The i.i.d. capacity loss due to quantization is also reported.

In the experimental part, verification of the concept in a laboratory environment was performed and study of the degradation effect on channel capacity due to the presence of PMD for an NRZ system was performed. The i.i.d. channel capacity loss due to: (i) PMD, (ii) A/D conversion, (iii) inadequate channel memory assumption, (iv) quantization of LLRs was determined and presented.

3.2 Capacity calculation

As explained in a previous paper [6] the optical channel is described by discrete dynamical trellis and is modeled as an ISI channel with memory \(2m+1\). Let \(X=(x_1, x_2, \ldots, x_n)\) and \(Y=(y_1, y_2, \ldots, y_n)\) denote the transmitted and the received sequences, respectively. The memory assumption means that a bit \(x_j\) is affected by the preceding \(m\) \((x_{j-m}, x_{j-m+1}, \ldots, x_{j-1})\) bits and the next \(m\) \((x_{j+1}, \ldots, x_{j+m})\) bits in the sequence. The channel is completely defined by \(X, Y\) and the conditional probability \(P(X|Y)\). Assuming that the transmitted sequence \(X\) is independent and uniformly distributed the information rate can be calculated from:

\[
I(Y;X) = H(Y) - H(Y|X).
\]

where \(H(U) = E(\log_2 P(U))\) is the entropy of a random variable \(U\) and \(E(\cdot)\) is the mathematical expectation operator. The Shannon-McMillan-Brieman [44] theorem states
that:

$$E\left( \log_2 P\left( Y^n_i \right) \right) = \lim_{n \to \infty} \frac{1}{n} \log_2 P(Y^n_i),$$

(3.2)

where $Y^n_i=(Y_1,\ldots,Y_n)$. Thus the information rate can be determined by calculating $\log_2(P(Y^n_i))$, for sufficiently long sequence. The first step for doing that is to modify equation (1). It can be rewritten as the summation

$$I(Y,X) = \lim_{n \to \infty} \frac{1}{n} \left[ \sum_{j=1}^{n} \log_2 P(y_j \mid y^{j-1}, x^n_i) - \sum_{j=1}^{n} \log_2 P(y_j \mid y^{j-1}) \right].$$

(3.3)

A state at discrete period in time $j$ is defined as $s_j=(x_{j-1},\ldots,x_j,x_{j+1},\ldots,x_{j+n})$. Let the set of all states be denoted as $S \{s_0, s_1, s_2,\ldots,s_k\}$, where $k=2^{2m+1}$ for binary transmission. The PDF is defined as $p(y_j \mid s)$ ($y_j$ is the sample that corresponds to $x_j$). The forward step of the BCJR algorithm calculates the probability of the terminal state $s_j$ based on the received sequence $(y_1,y_2,\ldots,y_j)$. That means that the forward metric is given by $a_j(s)=\log[p(s_j=s \mid y_1 \ldots y_j)]$. The conditional probability for the current received bit $y_j$ on the condition that all previous bits were received $(y_1,y_2,\ldots,y_{j-1})$ can be represented in the following manner:

$$P(y_j \mid y^{j-1}_i) = \sum_{j=1}^{n} \alpha_{j-1}(s_i) P_{il} P(y_j \mid s_j), \ i,l = 1,2,\ldots,k, j = 1,2,\ldots,n,$$

(3.4)

where $P_{il}$ is the probability of transition between the states from state $i$ to state $l$ from the previous time period $j-1$ to the present time period $j$. It is defined by the transition metric $\gamma_j(s_i,s_l) = \log[p(y_j \mid s_j)P(x_j)]$. Equation (3.4) is a summation over all possible transitions from a state $i$ to the observed state $l$ given the observed bit $y_j$. The meaning of the forward branch in this context is a weight distribution of the possible transitions that
lead to bit $y_j$ and state $s_l$. The forward metric can be recursively determined from all possible transitions that lead to the current state and the previous value of the forward metric as follows

$$
\alpha_j(s_i) = \max_{s_l} \left[ \alpha_{j-1}(s_i) + \gamma_j(s_i, s_l) \right]
$$

(3.5)

That means that the forward metric provides the weight of all transitions to a current state out of all possible transitions between any two given states. This can be expressed in the following equation.

$$
\alpha_j(s_i) = \frac{\sum_{i,j} \alpha_{j-1}(s_i) P_{il} P(y_j | s_l)}{\sum_{i,l} \alpha_{j-1}(s_i) P_{il} P(y_j | s_l)}, \quad i, l = 1, 2, \ldots, k
$$

(3.6)

From this expression it is clear that after initializing the forward metric, there are two sets of values left to determine before the mutual information can be computed. The probabilities of transition between states are equal for a uniformly distributed source of information. Then the last component we need to calculate are the conditional probabilities $P(y_j | s_l)$. To determine them we have used histograms of all states vs the received samples. This can be done by transmitting long enough sequence and creating a table with the desired resolution step for the samples.

### 3.3 Simulation results

The joint influence of different impairments is considered in this section. The first performed simulation was for a 10Gb/s RZ system operating in the presence of 100ps of digital group delay (DGD). The plot is shown in Fig. 3.1 and it shows the i.i.d. channel
capacity as a function of the optical signal to noise ratio (OSNR). The two curves serve to visualize the difference of performance for different memories of the BCJR equalizer. A memory of one bit period is sufficient for compensation of the impairment and full capacity is achieved, while memory of zero is not sufficient and the i.i.d. capacity of 0.5 is the maximum achievable value.

![Graph showing i.i.d. channel capacity simulations](image)

**Fig. 3.1** i.i.d. channel capacity simulations for DGD=100ps and 10Gb/s RZ system.

The second simulation was performed for a 10Gb/s NRZ system for 700km of SMF28 with 16ps/nm.km dispersion coefficient and dispersion slope of 0.08 ps/(nm².km). Duty cycle of 33% was used. The simulation was performed for state memories $m=0-3$. The results are presented in Fig. 3.2. It is clear that memory of $m=3$ for the BCJR equalizer is sufficient for compensation of 11200ps of chromatic dispersion.
Fig. 3. 2 i.i.d. channel capacity simulation for 10Gb/s NRZ system and 700 of SMF28.

Fig. 3. 3 i.i.d. channel capacity simulation for 10Gb/s NRZ system and 700 of SMF28 and 50ps of DGD.
The maximum achievable capacity gradually diminishes with the decrease of the memory. The capacity for memory zero is about 0.1, which cannot be compensated for by only equalization. To achieve maximum transmission capacity and utilize the bandwidth of the available fiber error correction coding has to be employed.

Fig. 3.3 shows the i.i.d. channel capacity in the system described in the previous simulation but in the presence of 50ps of DGD.

The added impairment deteriorates the values of the i.i.d. capacity compared to the previous case. Capacity of 1 is still achieved for memory of \( m = 3 \) in the BCJR turbo equalization scheme. However the observed penalty is about 4dB compared to the case where chromatic dispersion is the only impairment.

![Figure 3.4](image)

**Fig. 3.4** i.i.d. channel capacity as a function of the number of links/spans in a 40Gb/s system that consists of multiple links. The dispersion map is shown in Fig. 3.5.

The intra-channel four wave mixing is another major impairment for transmission at 40Gb/s and above. By increasing the memory of the turbo equalizer the transmission
distance of the system can be increased by up to 10 links or about 1200 km, as shown in Fig. 3.4. This is essential for the optimization of long-haul optical communication systems. The total transmission distance at full capacity for this simulation was 7200 km. To increase the distance even further, stronger codes need to be employed, the memory of the equalizer has to be increased, and the optimal dispersion map has to be employed.

Fig. 3.5 shows the dispersion map used in simulations in Fig. 3.4. The intra-channel four wave mixing is the dominant nonlinearity for this dispersion map. After the transmitter pre-compensation of -1600ps was used. N sections of 120 km consist of 80 km of D⁺ and 40 km of D⁻ fiber. D⁺ has attenuation constant 0.19dB/km, dispersion 20ps/nm.km, dispersion slope of 0.06 ps/(nm².km) and nonlinear refractive index of 2.6.10⁻²⁰ m²/W. D⁻ has attenuation constant 0.2dB/km, dispersion 40ps/nm.km, dispersion slope of 0.06 ps/(nm².km) and nonlinear refractive index of 2.6.10⁻²⁰ m²/W.

![Dispersion map for dominant intra-channel four wave mixing.](image)

Duty cycle of 33% was used. Launch power of 0dBm was used. The values for the optical and electrical filter were chosen as 3Rᵣ and 0.7Rᵣ respectively, where Rᵣ is the actual transmission rate calculated after dividing the desired data rate by the code rate such that the effective data rate is 40Gb/s.
Fig. 3.6 shows the effects of LLR quantization on the i.i.d. capacity for the same dispersion map and system configuration used in the previous simulation. It is apparent that there is a threshold of useful number of bits for LLR quantization. For the current setup 3 bits provide sufficient result and no significant improvement in capacity is observed for higher number of quantization bits. Significant channel capacity loss of 20% to 30% is found for 2 bit quantization.

![Fig. 3.6 i.i.d. channel capacity loss as a result of LLR quantization as a function of the number of spans for the dispersion map of Fig. 16 and RZ transmission at 40Gb/s.](image)

To visually demonstrate the effects the impairments have on the conditional PDFs, a simulation result is presented in Fig. 3.7. It illustrates the conditional PDFs of the received sample $y$ for two states: $s='11011'$ and '00100' and different values for DGD. The PDF mean for state $s='11011'$ shifts to the right, with the increase of DGD. Due to the insufficient memory of the PMD compensator, the residual ISI becomes more
significant as DGD increases. Clearly, the PDF curve becomes wider, and BER performance degrades when more DGD is introduced.

Fig. 3.7 Conditional PDFs for two states of the BCJR trellis in the presence of DGD.

3.4 Experimental results

Fig. 3.8 shows the setup used for this experiment. A pre-coded test pattern was uploaded into a pulse generator (Anritsu MP1763C) via personal computer (PC) with GPIB interface. The 10Gb/s NRZ signal was used to drive a zero chirp Mach-Zehnder modulator (UTP). Then the signal passed through PMD emulator (JDSU PE3), where controlled amount of DGD was introduced. The distorted signal was mixed with controlled amount of amplified spontaneous emission (ASE) noise with 3 dB coupler.
Modulated signal level was maintained at 0 dB while the ASE power level was changed to obtain different OSNRs. Next, the optical signal was pre-amplified (Optigain, Inc. 2000 series), filtered (JSDU 2nm band-pass filter) followed by detection (Agilent 11982A). An oscilloscope (Agilent DCA 86105A) triggered by the data pattern was used to acquire the samples. To maintain constant power of -6 dBm at the detector, a variable attenuator was used. Data was transferred via GPIB back to the PC, which employed the forward step of the BCJR algorithm to calculate channel capacity.

Fig. 3.9 shows the experimental results for the i.i.d. capacity for different values of the trellis memory ($2m+1=5$ and $2m+1=1$).
Fig. 3. 9 Memory effects on i.i.d. channel capacity: (a) BER plots at memory $2m+1=5$, (b) BER plots at memory $2m+1=1$. 
Unity capacity is achievable in both cases although the case of memory 0 and DGD of 1.25 indicates that for higher values of DGD capacity of 1 will be impossible to achieve. It is clear that the usage of dynamic discrete model for the optical channel is needed to effectively reduce PMD effects.

Fig. 3.10 shows how the quantization effects influence the i.i.d. channel capacity loss. The Fig. 3.10 (a) studies the quantization of the LLRs, while the Fig. 3.10 (b) studies the influence of A/D conversion on capacity. Let the real LLR be denoted by $\lambda$. Calculations were done with double precision numbers. A 2-tuple of integers $(d_b, p_b)$ was used to represent the integer (with sign) and decimal part of $\lambda$, respectively [45]. For example, in $(5,0)$ representation, 5 denotes the number of bits used to represent the integer part of $\lambda$ and 0 represents the number of bits used for decimal part. The quantization step size is determined by the maximum and the minimum sample value and the number of bins needed for quantization. Fig. 3.10 (a) demonstrates that $(5,0)$ and $(4,0)$ representations are very close to the double precision one. Significant i.i.d. channel capacity degradation is observed for the $(3,0)$ case. Compared to the simulation results where 3 bit quantization for LLRs was sufficient, in an experimental environment 4 bits are required.

From Fig. 3.10 (b) we conclude that the i.i.d. channel capacity is less sensitive to A/D quantization compared to LLR quantization. The degradation here is more gradual compared to the case of LLR quantization.
Fig. 3. 10 Quantization effects influence on i.i.d. channel capacity (C): (a) LLR quantization, (b) A/D conversion.
Significant loss is noticed for the (3,0) case and is not that much severe as in the LLR case. For number of quantization bits 4 and above there is improvement but it is not very significant. We can conclude that 4 bits is the optimal number of quantization bits for both LLR and A/D.

3.5 Conclusion

In conclusion, we proposed a channel capacity calculation method based on the forward step of the BCJR algorithm and conditional PDFs determination and verified by simulations and experimental demonstration. The effects of chromatic dispersion, PMD, intra-channel four wave mixing on channel capacity were studied and reported. The effects of BCJR memory, LLR quantization and A/D quantization on the channel capacity were investigated and reported.

It was demonstrated that the proposed method is capable of accurate capacity calculation. Capacity of 1 is achievable for all configurations provided enough memory for the BCJR equalizer can be used. Memory of $m=3$ was sufficient for compensation of 11200ps of DGD for 10Gb/s NRZ transmission. For RZ modulation format and 40Gb/s transmission rate in the presence of intra-channel four wave mixing, memory $m=3$ was sufficient to fully compensate the impairment for up to 4000km with transmission capacity of 1, provide 90% capacity for up to 6000km and 80% for up to 7200km. Higher memories are expected to compensate for even longer transmission distances. The memory of $m=3$ was sufficient to compensate for simultaneous impairments of 50ps DGD and 11200ps of chromatic dispersion and achieve channel capacity of 1.
For NRZ transmission at 10Gb/s DGD of up to 125ps was successfully compensated for to achieve channel capacity of 1 for memory of \( m=3 \). It is expected that higher values of DGD can be compensated for to achieve maximum capacity as simulations indicate.

Simulations and experimental results agree on the existence of optimal value for the number of quantization bits for LLRs and A/D conversion. The experimental results resulted in 4 bits for LLRs and 3 bits for A/D as optimal value while the simulations indicated 3 bits for the LLRs as sufficient.
CHAPTER IV - LDPC CODES FOR HIGH SPEED OPTICAL COMMUNICATION SYSTEMS

4.1 Introduction

Forward error correction (FEC) for optical communication systems is an active research topic in recent years. The latest developments in this area have lead to the rapid increase of data rates at which these systems operate. However with the higher data rates new problems have emerged. At transmission speeds of 40Gb/s and above, fiber nonlinearities are the dominant signal impairment. One of the primary methods for mitigation of signal impairments is the use of error correction coding. Such techniques have been successfully used in fiber-optic transmission systems for years. However previous generations of optical systems have been operating at lower speeds at which fiber nonlinearities are not the dominant impairment. This has lead to the necessity of improved techniques for mitigation of the effects of nonlinear impairments on the signal. Error correction coding can be successfully used to mitigate both linear and nonlinear effects of the propagating signal but new, stronger codes are highly desirable. The increased importance of the coding techniques in optical communication systems, and the
fact that the performance limits of hard decision decoding systems have almost been 
exhausted, has lead to the intensive research of soft decoding algorithms which perform 
closer to the theoretically achievable performance.

Codes that employ soft-decision iterative decoding have emerged as promising 
candidates for replacement of the currently standardized concatenated Bose-Chaudhuri-
Hocquenghem/Reed-Solomon (BCH/RS) codes. Soft decision iterative decoding has 
been proven to reach the Shannon capacity limits in real systems for sufficient length of 
the code. Among the codes that have been researched in recent years there are two major 
candidates that are considered viable for use in future generations of optical 
communication systems. The first type of codes are turbo product codes with BCH 
component codes [14]. The second type of codes are low density parity check (LDPC) 
codes [7],[45][30],[31],[45]-[48]. The LDPC codes can be classified in two categories 
according to their structure. There are random codes and structured codes. Good random 
LDPC codes provide excellent coding gains due to the random structure of their parity 
check matrices. However, the decoding complexity associated with these codes is beyond 
the capabilities of modern hardware. With equal other parameters the random codes have 
better performance compared to the systematic codes. Regardless of that fact they are not 
only very computationally heavy to decode but also more difficult to generate in a 
systematic fashion. It has been previously shown that structured LDPC codes are able to 
match and outperform turbo-product codes in terms of coding gain and decoder 
complexity [45]. It has also been shown that structured codes can approach the 
performance of random codes and even outperform them in some cases [49]. The benefits
of easy generation, excellent performance and decoding complexity low enough to be implemented in hardware [7],[45],[30],[31],[45]-[48] have made the structured LDPC codes most attractive candidate for the near future optical communication systems.

In order to improve the performance of the structured codes and compensate for the performance loss due to the regularity of the parity-check matrices, different code parameters can be optimized. The parameters that can be controlled are the column-weight of the parity-check matrix, the code-word length, the girth (the shortest cycle length in the corresponding Tanner (bipartite) graph representation of the parity-check matrix), and trapping sets.

Among all the parameters for increasing the code gain, the girth of the code is the most effective and reliable one. It indicates the shortest cycle in the graphical representation of the code’s parity-check matrix. Higher girth codes perform significantly better. This is related to the decoding of the codes as explained in earlier chapter. Increasing the code girth also doesn’t increase the complexity of the decoder at all, which is essential to the implementation of the codes in hardware.

Previous simulations of large girth LDPC codes have shown outstanding results and coding gains. In this chapter we have experimentally tested the performance of several large-girth codes with varying other parameters – length, girth and construction type. The two structured codes that have been explored are the quasi-cyclic codes also known as array or block-circulant and balanced incomplete block design (BIBD).

The length of the code is a limiting factor to the large girth as it has been shown in [31]. There is minimum length of the code required for a certain girth. The maximum
girth for quasi-cyclic codes is 12 and the girth has to be an even number by definition. The reason for that is that higher girths would make it impossible to preserve the quasi-cyclic structure of the parity check matrix of the code. Codes of girth 6 are relatively easy to design. In this chapter, codes of girths 8 and 10 are considered. There are different approaches to the design of such codes. The codes that have been used in this chapter have been designed with an efficient computer search algorithm. This imposes a second limitation on the length of the codes. Further to evaluate the performance of these codes in real conditions we studied them for use in a fiber optic communication system operating at 10Gb/s. However, the LDPC codes require soft information for decoding and that renders most commercially employed equalizers useless for this purpose. Thus the equalization scheme of choice for this study was a turbo-equalization scheme which uses soft decision information. The description of the turbo equalization scheme is given in the next chapter. They are iteratively providing information to each other resulting in best overall decoding performance.

### 4.2 BIBD code design

One of the first attempts at structured design for LDPC codes has been provided by [50]. The balanced incomplete block design is a combinatorial structure [51],[52]. Here we are concerned with design of LDPC codes suitable for use in optical communication systems and their experimental performance evaluation. The balanced incomplete block is a collection of \( b \) subsets of a given set \( V = \{x_1, x_2, \ldots, x_v\} \) called blocks. The elements of \( V \) are called points and their number is denoted by \( v \). Let the blocks are denoted by \( B_1, B_1, \ldots, B_m \). Let there be an integer parameter \( t \) such that \( t \leq v \). If the following conditions
are satisfied, a BIBD is of \( t \)-design type: 1) any subset of \( t \) points is contained in the exactly the same number of blocks \( \lambda \) (or \( t \)-design); 2) each block contains the same number of points \( k \). There are three parameters for the BIBDs that can be considered as independent. This is given by the relationship among the parameters. Equations (1) and (2) define these relationships.

\[
vr = bk \\
\lambda(v-1) = r(k-1)
\]

The notation used for these codes is \((v,k,\lambda)\)-BIBD. In the current chapter only codes with \( t=2 \) are considered. We define any BIBD code design to be resolvable if there exists a nontrivial partition of its blocks set \( B \) into parallel classes such that each class partitions the set of points \( V \).

\[
H = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 
\end{bmatrix}
\]

The parity check matrix above is an example of BIBD design. For that specific example we look at the following collection of blocks: \( B = P\{\{1,2,3\},\{4,5,6\},\{7,8,9\},\{1,4,7\},\{2,5,8\},\{3,6,9\},\{1,5,9\},\{2,6,7\},\{3,4,8\},\{1,6,8\},\{2,4,9\},\{3,5,7\}\} \) is a 2-design with parameters \( v=9, b=12, \) and \( \lambda=1 \). This design is resolvable with four resolvability classes each of which contains three blocks. If represented in a graphical way the
resolvability classes would be lines connecting the points of the code. The blocks of the code are also related to the parity check matrix of the code. If the points are considered as the parity-check equations and the blocks are considered as the bits of a linear block code the parity-check matrix of an LDPC code is defined.

The rate of the code constructed in such a way is $R=(b\text{-rank}(H))/b$. Several types of systematic constructions exist for BIBD codes. The code used in the current research is based on the theory of orthogonal arrays [53], and references therein]. An orthogonal array of size $N$ and constraint $k$, $q$ levels, strength $t$, and index $\lambda$, is denoted as $\lambda$-OA($N,k,q,t$). It is defined as an $k\times N$ matrix $A$ with entries from a set of $q$ elements such that any $t\times 1$ column vector in $t\times N$ submatrix is contained $\lambda$ times. Any $t$ rows of the $\lambda$-OA($N,k,q,t$) matrix contain each ordered pair of elements exactly once. The orthogonal arrays of strength 2 and index unity are denoted by OA($N,k,q$) and it had been shown [54],[56] that provided the condition $N=q^2$ the problem of existence of OA is equivalent to the existence of $k$-2 mutually orthogonal Latin squares (MOLSs). The mutually orthogonal Latin squares are a subclass of the mutually orthogonal Latin rectangles (MOLRs). The MOLR construction of BIBD codes has the property of flexible number of blocks $b$ and arbitrary block size $k$ which results in high-rate and low length of the codes which is suitable for optical communication systems.

For the construction to be resolvable the MOLRs have to satisfy certain conditions. The Latin rectangle is defined as an $m\times k$ array $[L_k(x,y)]_{1\leq x\leq k,1\leq y\leq m}$ with elements in $L_k(x,y)$ in $\{0,1,\ldots,m\}$ provided that each of the elements occurs once in each row and once in each column. Next a join of two rectangles
$L_k^1 = \left[ L_k^1(x,y) \right]_{1 \leq k \leq m}$ and $L_k^2 = \left[ L_k^2(x,y) \right]_{1 \leq k \leq m}$ is defined as an $m \times k$ array whose $(x,y)$ entry is the pair $(L_k^1(x,y), L_k^2(x,y))$ and is denoted by $(L_k^1, L_k^2)$. If the two Latin rectangles are orthogonal all their entries in the join are distinct. Latin rectangles $(L_k^1, L_k^2, \ldots, L_k^n)$ are mutually orthogonal if they are orthogonal in pairs. When the dimensions of the rectangles are equal and a power of a prime number we can construct a BIBD LDPC code in the following manner. Let $k = m = p^l$, where $p$ is a prime number and $l \geq 1$. Let the elements of $GF(k)$ are $\alpha_0 = 0, \alpha_1, \alpha_2, \ldots, \alpha_{k-1}$. Then a $k \times k$ array can be defined with the non-zero element $\alpha$: $L_k^\alpha(x,y) = \alpha \alpha_x + \alpha_y$, $0 \leq x, y \leq k - 1$. Designed in this fashion the MOLRs are mutually orthogonal of order $k$ [54]. Next an integer lattice is considered which elements are labeled by the points of the set $V$. Let $l$ is a one-to-one mapping of the square $L$ is to the integer set $V$, $l : L \to V$. The numbers $l(x,y)$ are called cell labels. Every rectangle $L_k^\alpha$ defines a set of $k$ parallel lines in the integer lattice. Those lines can be written as $B_k^\alpha = \{ L_\alpha(x,y) = s \}, 0 \leq s \leq k - 1$ and they are equivalent to the set of lines with $0 \leq s \leq k - 1$. For a given $GF(k)$ the set of parallel lines define the positions of 1s in the parity check matrix of the BIBD code. Every set of lines defines a submatrix within the parity check matrix.

$$H_{km,km} = \begin{bmatrix}
H_{1,1} & H_{1,2} & H_{1,3} & \cdots & H_{1,m} \\
H_{2,1} & H_{2,2} & H_{2,3} & \cdots & H_{2,m} \\
H_{3,1} & H_{3,2} & H_{3,3} & \cdots & H_{3,m} \\
\vdots & \vdots & \vdots & \cdots & \vdots \\
H_{k,1} & H_{k,2} & H_{k,3} & \cdots & H_{k,m} 
\end{bmatrix}$$ (4.3)
To increase the code girth, a computer algorithm [55] has been designed, which ensures the minimum girth during construction. By employing this algorithm the code LDPC(8547,6922), R=0.81, g=8, w=4 was designed. This code was used for the study in the current chapter.

4.3 Large girth quasi-cyclic LDPC codes

The majority of codes under study in this chapter belong to the quasi-cyclic (QC) class of LDPC codes. These codes are based on the theory of generalized difference sets and more precisely on \( m+1 \)-fold cycle-invariant difference sets (CIDSs) introduced in [57]. The general difference set \((v,k,\lambda)\) is defined as a set \( S = \{s_1, s_2, \ldots, s_k\} \) over an additive Abelian group \( V \), such that every non-zero element \( s \) from \( V \) can be represented as a subtraction of two elements of \( S \), \( s = s_i - s_j \) in at most \( \lambda \) ways. For the applications of optical communication systems high-girth codes are desirable and it has been shown that for \( \lambda \geq 2 \) the generated codes have low girths. That is why only \( \lambda = 1 \) codes are considered in this study.

To define the \( m+1 \)-fold cycle-invariant difference set a cyclic shift operator is used over an additive group \( Z_N \) of integers modulo \( N \). Let the elements of a general difference set \( S \) over \( Z_N \) be arranged in a given order. Let the cyclic-shift operator that cyclically shifts a sequence \( t \) positions to the right is denoted by \( C^t \). If the ordered sets \( R_i = C^i S - S \mod N \) are themselves difference sets over \( Z_N \) modulo \( N \), the set \( S \) is defined as an \( m+1 \)-fold cycle-invariant difference set. The subtraction is performed component-wise. Now the parity-check matrix of the block-circulant LDPC codes based
on CIDSs can be constructed in the following manner

\[
H = \begin{bmatrix}
p_{i_1} & p_{i_2} & p_{i_3} & \cdots & p_{i_s} \\
p_{i_1} & p_{i_2} & p_{i_3} & \cdots & p_{i_{s-1}} \\
p_{i_{s-1}} & p_{i_s} & p_{i_1} & \cdots & p_{i_{s-2}} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
p_{i_{s-m+2}} & p_{i_{s-m+3}} & p_{i_{s-m+4}} & \cdots & p_{i_{s-m+1}}
\end{bmatrix},
\]

(4.4)

where \(i_1, i_2, \ldots, i_s\) are non-negative integers and \(P\) is a \(N \times N\) basic circulant permutation matrix. The dimension of the resulting parity check matrix is \((mN) \times (sN)\) and the elements of \(H\) are the powers of the basic permutation matrix \(P\).

\(P\) is given by

\[
P = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \ddots & \cdots \\
0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

The row weight of the code given by \(H\) is \(s\) and the column weight is \(m\). The code given by \(H\) is an intersection of quasi-circulant or quasi-cyclic code of order \(s\). A quasi-cyclic code or order \(s\) is such that any cyclic shift of a codeword with \(s\) positions results in another codeword. Intersection of a class of codes \(C\) is the sets of all codewords that belong to every code in the class.

A simplified version of (2) exists which results in a lower number of cycles within the matrix. It can be obtained from (2) when the powers of the permutation matrices from the first row and column are set to 0. This form of the parity check matrix is useful for generating high-girth LDPC codes.
where $I$ is $p \times p$ (p is a prime number) identity matrix, $D$ is $p \times p$ permutation matrix ($d_{i,i+1}=d_{p,i+1}=1$, $i=1,2,\ldots,p-1$; other elements of $D$ are zeros), while $r$ and $c$ represent the number of rows and columns in (1), respectively. The set of integers $S$ are to be carefully chosen from the set $\{0,1,\ldots,p-1\}$ so that the cycles of short length, in the corresponding Tanner graph representation of (1), are avoided. We have shown in [48] that large girth, $g=10$, LDPC codes provide excellent improvement in coding gain over corresponding turbo-product codes (TPCs). At the same time the complexity of LDPC codes is lower than that of TPCs, selecting them as excellent candidates for application to systems for beyond 40 Gb/s transmission. Namely, the minimum distance for an LDPC code is given by the Tanner bound [27]:

$$
d \geq \begin{cases} 
1 + \frac{w}{w-2} \left( (w-1)^{(g-2)/4} - 1 \right), & g / 2 = 2m + 1 \\
1 + \frac{w}{w-2} \left( (w-1)^{(g-2)/4} - 1 \right) + (w-1)^{(g-2)/4}, & g / 2 = 2m 
\end{cases}
$$

(4.6)

where $g$ and $w$ are respectively the girth of the LDPC code and the column weight of the parity check matrix. The operator $\lfloor \rfloor$ indicates the largest integer that is smaller or equal to the enclosed number. Equation (4) shows that the linear increase in the girth results in exponential increase of the minimum distance. Notice that this bound is tight only for short codes (in the order of hundreds), nevertheless it provides a guideline of how to design the LDPC codes of large minimum distance. For example, by selecting $p=1123$.
and $S = \{2, 5, 13, 20, 37, 58, 91, 135, 160, 220, 292, 354, 712, 830\}$ an LDPC code of rate 0.8, girth $g = 10$, column weight 3 and length $N = 16845$ is obtained. This code and the LDPC codes LDPC(11936,10819), $R = 0.906$, $g = 10$, $w = 3$ and LDPC(15328,13893), $R = 0.906$, $g = 10$, $w = 3$ are used in this study.

### 4.4 Generalized LDPC (GLDPC) code design

The generalized LDPC codes are first introduced by Tanner [27]. They provide benefits of excellent minimum distance and low-complexity decoding [58],[59]. When properly designed their performance approaches Shannon’s limit. The construction of GLDPC codes consists of replacing the parity-check equation of an LDPC code with the parity-check matrix of a simple linear block code. The LDPC code is known as a global code and the simple linear block codes used for substitution is known as a constituent or local code. Usually the local codes are Hamming, BCH or Reed-Muller codes. Different constructions of GLDPC codes have been proposed [33],[34],[56],[60]-[62]. The decoding of GLDPC codes is based on the implementation of a simple and fast soft-input-soft-output (SISO) decoding for the short local linear block codes. Usually MAP decoding is implemented for the SISO decoders. For Hamming and Reed–Muller codes a simple MAP algorithm based on the modified Walsh–Hadamard transform proposed in [59] can be utilized. In this study we have investigated GLDPC codes for which the local codes are Hamming codes. The decoding on the local level was accomplished through the use of an appropriate number of BCJR decoders operating in parallel.
There are three types of GLDPC codes based on the component codes used: 1) GLDPC codes with algebraic local codes of short length – Hamming, Reed-Muller; 2) GLDPC codes with high-rate LDPC local codes with large minimum distance; 3) GLDPC codes with GLDPC local codes, or so called fractal GLDPC codes. Each of these provides different specific advantages. The first construction utilizes MAP decoding for the local codes which results in small number of decoding iteration and accurate estimates for the LPDC variables. This results in very small number of iterations (typically below 10) over the LDPC graph to achieve very low BERs. The second construction provides good code designs of high codeword length. The fact that both the global and the local codes are LDPC codes with very high minimum distances, results in an even higher aggregate minimum distance, which significantly enhances the code performance. The third construction offers attractive possibilities for implementation in very large scale integration (VLSI) designs due to its scalability and the fact that the same structure can be used on different levels.

In this study we have limited our choice to a single representative of GLDPC code with construction of the first kind with Hamming local code. The construction of that code is described next. The other constructions are described by [34]. Calculation of the lower bound on the minimum distance of such GLDPC codes is given in [33].

The parity check matrix of a code is partitioned into $W$ submatrices in the following manner $H = \begin{bmatrix} H_1^T & H_2^T & \cdots & H_w^T \end{bmatrix}^T$, where every submatrix is a block-diagonal matrix which is constructed by replacing the 1s in the main diagonal of an identity matrix with
the parity-check matrix $H_0$ of the local code. The local code is denoted by $C_0(n,k)$. Each of the matrices $H_2,..., H_W$ is derived from $H_1$ by performing a random column permutation $\pi_{j-1}, j=2,\ldots,W$.

$$H_1 = \begin{bmatrix} H_0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & H_0 \end{bmatrix}, \quad H_j = \pi_{j-1}(H_1), j = 2,\ldots,W$$  \hspace{1cm} (4.7)

The rate $R$ of the resulting GLDPC code is determined by the rate of the local code $r = k / n$.

$$R = \frac{K}{N} \geq 1 - W \left(1 - \frac{k}{n}\right),$$  \hspace{1cm} (4.8)

where $K$ and $N$ are the dimension and the code length for the GLDPC code. A GLDPC code constructed in this manner will be denoted as GLDPC$(N,W,n)$. It is the intersection of $W$ super-codes $C_1, C_2,\ldots, C_W$ with parity-check matrices $H_1, H_2,\ldots, H_W$, respectively. The parameter $W$ determines the column-weight of the GLDPC code. It has been shown in [60] and [61] that $W=2$ leads to excellent BER performance.

The decoding for this type of GLDPC construction can be performed in an iterative manner as described in [58][60][61]. Every received codeword consists of $N$ samples. On the first step of decoding every received bit is assumed to belong to the super code $C_1$ with parity check matrix $H_1$. The reliability for this bit is calculated from the acquired sample by using $N/n$ SISO decoders that operate in parallel on the independent codes $C_0$ that comprise the code $C_1$. The SISO decoders calculate a posteriori probability and extrinsic probability for every coded bit. Once the first step of decoding is complete, the
procedure is repeated but this time for the super code \(C_2\). It is a second level of nested decoding. The inputs to the second step are the computed extrinsic probabilities for every bit on the previous step. Again they are \(N\) values separated into \(N/n\) sections fed into \(N/n\) parallel SISO decoders with the parity check matrix \(H_0\). Because code \(C_2\) is a permutation of code \(C_1\) the same set of decoders as the first step can be used but the input vector is permuted according to the permutation that transforms \(C_1\) into \(C_2\). On this step extrinsic probabilities of code \(C_2\) are calculated again fed into \(N/n\) constituent codes \(C_0\) of super-code \(C_3\). On the final step of decoding the extrinsic probabilities of super-code \(C_{W,1}\) are fed into \(N/n\) constituent codes \(C_0\) of super-code \(C_W\). This completes the first iteration of decoding. A terminal number of iterations are performed or the procedure is stopped if a valid codeword has been reached. The SISO decoders can be implemented as \(N/n\) identical parallel low-complexity BCJR decoders for every \(H_0\). Additional decrease of the decoding complexity can be achieved by using simplified version of the BCJR algorithm or by employing a simple MAP decoding algorithm proposed in [62]. For GLDPC codes with \(W=2\) the GLPC code can be considered as a generalization of a turbo decoder where two BCJR algorithms are used for every super-code [27]. To increase the girth of the GLPC codes the cycle removal method in [63] can be employed. In this fashion the code GLDPC(6048, 1152) \(R=0.8095, g=8, w=2\) was created and used in the current study. It employed the general code LDPC(6048, 192) \(R=0.9683, g=8, w=3\) and the local code Hamming(63,6), \(R=0.90476\).
4.5 Experimental setup and results

The experimental setup used in this paper is shown in Fig. 4.1. A personal computer via GPIB interface controls the pattern generator (Anritsu MP1763C) for pattern upload and the oscilloscope (Agilent 11982A) for data collection.

Fig. 4.1 Experimental setup.

A zero-chirp Mach-Zehnder modulator (JDSU OC-192) is used for optical modulation of 10Gb/s NRZ electrical signal. Controlled amount of first-order PMD and amplified-spontaneous emission (ASE) noise are introduced with the aid of a PMD emulator (JDSU PE3), and an ASE source. Amplification is done with an erbium doped amplifier (Optigain, Inc. 2000 series), and photo-detection is performed with a detector Agilent 11982A. Constant level of 0dBm is maintained at the laser source output and the optical signal to noise ratio (OSNR) is controlled by a variable optical attenuator (VOA) and monitored with an optical spectrum analyzer. Constant level of -6dBm is maintained at the detector with the second VOA. The oscilloscope is triggered by the pattern
generator. The personal computer serves as an LDPC decoder (the hardware is currently in development stage).

The performance of several large-girth quasi-cyclic LDPC codes was experimentally evaluated. The LDPC codes under study are listed in Table 1.

<table>
<thead>
<tr>
<th>LDPC code</th>
<th>Code length</th>
<th>Code rate, $R$</th>
<th>Code type</th>
<th>The girth, $g$</th>
<th>Column weight, $w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDPC1</td>
<td>8020</td>
<td>0.950</td>
<td>BIBD</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>LDPC2</td>
<td>11936</td>
<td>0.906</td>
<td>QC</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>LDPC3</td>
<td>15328</td>
<td>0.906</td>
<td>QC</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>LDPC4</td>
<td>16935</td>
<td>0.800</td>
<td>QC</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>LDPC5</td>
<td>4376</td>
<td>0.936</td>
<td>MacKay</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>LDPC6</td>
<td>8547</td>
<td>0.81</td>
<td>BIBD</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>LDPC7</td>
<td>7225</td>
<td>0.93</td>
<td>OA product</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>GLDPC</td>
<td>6048</td>
<td>0.81</td>
<td>QC+H</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

For comparison purposes the random LDPC5 code (MacKay’s random code) is evaluated as well. To guarantee reliable calculation sequence of length approximately 64000 bits is used and acquired multiple times resulting in minimum 10 million test bits per measurement.
Fig. 4.2 Performance evaluation of LDPC codes without impairments for different values of OSNR.

Fig. 4.2 shows the results for the different codes that have been tested. The code performance down to BER=$10^{-6}$ has been tested. Three groups of codes can be distinguished according to coding gain. Group 1 has coding gain in the order of 4.5 dB, group 2 in the order of 7 dB and group 3 is 8 dB coding gain. The first group of codes includes the random code LDPC1 and code LDPC7. The second group includes codes LDPC4, LDPC2 and the GLDPC code with Hamming local code. The third group is formed by the code LDPC4.

The random code outperforms code LDPC7 by 0.25 dB at BER=$10^{-6}$. The random
code performs identical to code LDPC1 down to BER=$10^{-6}$. Code LDPC4 outperforms code LDPC2 by 1dB at BER=$10^{-6}$. GLDPC code with Hamming local code, code LDPC2 and code LDPC6 perform the same at BER=$10^{-6}$.

4.6 Conclusion

In conclusion different designs for LDPC codes have been implemented and their performance have been experimentally verified. We have shown that the structured LDPC codes can outperform the random construction LDPC codes by varying different parameters. The most important parameter is the code girth. Code LDPC4 performs the codes from the second group by 1dB at BER=$10^{-6}$. This includes code LDPC2 which has comparable girth and length but higher rate; code LDPC6 which has the same rate but lower girth and lower length; and GLDPC code with Hamming local code which has the same rate but lower girth and lower length.

The structured LDPC have been proven to provide excellent coding gains – up to 8dB at BER=$10^{-6}$. Even higher coding gains are expected for codes of higher girths and lengths. This makes them suitable for high-speed long-haul optical communication systems because they provide the low decoding complexity required for hardware decoder implementation, and allow for lowering the overhead without additional penalty to BER performance.
CHAPTER V - TURBO EQUALIZATION

5.1 Introduction

Currently wavelength division multiplexing (WDM) is the most common way of increasing capacity in optical communication systems. The bandwidth demand has increased rapidly in recent years leading to increase in transmission rates and smaller channel separation in WDM systems. At higher transmission rates WDM system performance is significantly degraded primarily due to fiber nonlinearities and polarization mode dispersion (PMD). Different digital signal processing techniques [5],[64]-[66] and equalization techniques [4],[6],[67] have been proposed with regard to this problem. The main benefits of electronic compensation are the low cost of the devices, the ability to use parallel processing and the integration possibilities on the receiver side.

Turbo equalization schemes have emerged as one of the major candidates for PMD compensation. The type of error correction codes are related to the choice of equalization scheme. For example the turbo equalization scheme based on convolutional codes proposed in [67] experiences an error floor around BER=10^-6 which is too high for the requirements of the telecommunication industry (BER in the range of 10^-15). In the past 10 years LDPC codes have gained significant importance in the area of optical communications. Special LDPC codes have been design to fit the needs of optical transmission systems [30][34][48]. They are structured and this results in very low
decoding complexity. It has been shown that structured LDPC codes have lower complexity compared to concatenated turbo product codes and they also outperform them in BER performance [9]. It had been shown that LDPC codes do not exhibit error floor down to BER of $10^{-15}$ [2]. That makes LDPC codes an outstanding candidate for suppression of PMD and fiber nonlinearities compensation in optical transmission systems.

As explained in a previous chapter LDPC decoding is using bit reliabilities and thus a hard decision equalizer cannot be implemented efficiently with these codes. The Bahl–Cocke–Jelinek–Raviv (BCJR) equalizer is a maximum a posteriori (MAP) decoder that provides soft decision information (log likelihood bit ratios) in addition to the detected sequence. By providing soft decision information such an equalization scheme enables the full potential of LDPC codes and an iterative turbo equalization scheme has been proposed as a result. The compensation of chromatic dispersion based on this scheme has been simulated [68], and compensation of PMD has been simulated and experimentally demonstrated [7],[69]. However in a real system no single impairment will be completely dominant and it has been shown that the proposed scheme is capable of simultaneous mitigation of multiple impairments [6],[70].

The proposed scheme consists of a BCJR equalizer and LDPC decoder that exchange information in an iterative fashion. The BCJR serves as a maximum \textit{a posteriori decoder} and lowers the BER down to the FEC threshold at which LDPC codes can perform efficiently. It partially cancels intersymbol interference (ISI), calculates
initial reliabilities for the LDPC decoder and accepts extrinsic information from the decoder. The iterative process of passing extrinsic information is referred to as turbo equalization. Several parameters are of importance in this scheme. The most sensitive optimization parameter for the LDPC code performance is the code girth. Large girth increases the code minimum distance and the BER performance of the code and decorrelates the extrinsic information passed to the BCJR equalizer by the LDPC decoder. To appropriately match the performance of the LDPC codes to the BCJR equalizer, extrinsic information transfer (EXIT) charts are employed [71]. This is an essential step of the scheme optimization. The decoding of LDPC codes has been explained in the previous chapter.

5.2 BCJR equalizer

Memoryless channel model is not appropriate for optical transmission systems due to the presence of nonlinear impairments such as PMD, inter- and intra-fiber nonlinearities, which cause nonlinear ISI. This is the reason for assuming a more adequate nonlinear ISI channel with memory $2m+1$. The channel model is discrete optical representation of the optical channel. The BCJR algorithm operates on this model. Based on a received vector of samples or received sequence $Y$, it provides the maximum a posteriori probability of the transmitted sequence or transmitted vector of bits $X$. Let at discrete period of time $j$ the received sample is $y_j$ and it corresponds to the transmitted symbol $x_j$ for that discrete time period. Because a memory assumption was made, the bit $x_j$ is influenced by the preceding $m$ $(x_j-m, x_j-m+1, \ldots, x_j-1)$ bits and the next $m$ $(x_j+1, \ldots, x_j+m)$ bits in the sequence. The BCJR algorithm operates on the discrete trellis at every discrete
period in time and memorizes the computations made for the \( m \) bits ahead and \( m \) bits behind. The collection of \( 2m+1 \) bits around the current bit is defined as a state \( s=(x_i, m, \ldots, x_i, x_{i+1}, \ldots, x_{i+m}) \). Different discrete periods of time correspond to a different state for that bit thus uniquely associating every received bit with one of \( 2^{2m+1} \) states. The trellis on which the BCJR algorithm operates consists of all possible states at all possible discrete periods of time. The algorithm calculates all possible transitions for every time an makes a decision about the most probable transmitted bit for every discrete period of time. To uniquely define the trellis in time the triple \{previous state, channel output, next state\} is used. Fig. 5.1 illustrates the trellis for memory \( 2m+1=5 \), at time instant \( j \). The preceding time instance is shown but the succeeding time instance is not represented in the diagram. The trellis has \( 2^{2m+1}=32 \) states, each corresponding to the different possible 5-bit patterns.

![Trellis Diagram](image)

**Fig. 5.1** Example of channel description with trellis the time periods \( j \) and \( j-1 \).

Let \( S \) denote the set of all possible states \( \{s_0, s_1, s_2, \ldots, s_{32}\} \). From equalizer perspective the bit that has to be detected for the current time period is the middle bit of the state corresponding to that time period. On the figure that is the middle bit from every
5 bit sequence next to a state. The arrows connecting the states are called edges and they indicate possible transitions from a given discrete period of time to the next one. There are two labels above every edge. The first one indicates the transmitted bit and the second one the bit in terminal state. For the equalizer to be fully functional and the optical channel to be fully characterized, the conditional density functions $p(y_j|x_j)=p(y_j|s)$, $s \in S$ are needed. They are determined by the collecting the histograms by pre-transmitting a sufficiently long training sequence.

The weights of every possible transition enable the equalizer to determine the most probable path in the trellis. They are called metrics. The algorithm makes two passes for every bit sequence – one in the forward direction and one in the backward. That is needed because both precursor and postcursor ISI are taken into account. The metric calculation rule is shown in Fig. 5.2.

The forward metric indicates the probability of the terminal state being $s$ given a received sequence $(y_1, y_2, \ldots, y_j)$. In log-domain the corresponding reliability is given by $\alpha_j(s)=\log[p(s_j=s, [y_1 \ldots y_j])]$, $(j=1,2,\ldots,n)$. The forward metric is propagating from previous moment of time to present moment of time (i.e. propagates forward in time). The value of the present forward metric is recursively dependent on the value of the previous forward metric. The backward metric is related to the conditional probability of receiving the next $n-j$ bits given the current state. In log-domain the reliability is defined by $\beta_j(s)=\log[p([y_{j+1} \ldots y_n]|s_j=s)]$. The backward metric is propagating from present moment of time to previous moment of time (i.e. propagates backward in time).
Fig. 5.2 Forward and backward recursions for the BCJR equalizer.

The value of the previous backward metric is recursively dependent on the value of the present backward metric. The branch metric is defined as the transition probability between two states when a bit $y_j$ has been received. The corresponding reliability in log-domain is defined by $\gamma_j(s',s) = \log[p(s_j=s, y_j=s_j, a_{j-1}=s')]$. The branch metric is dependent on the present values of the forward metric and the backward metric. The metrics have to be updated iteratively for every discrete period of time as given by the update rules:

$$\alpha_j(s) = \max_{s'} \left[ \alpha_{j-1}(s') + \gamma_j(s', s) \right]$$

(5.1)

$$\beta_{j-1}(s') = \max_{s} \left[ \beta_j(s_1) + \gamma_j(s', s_1) \right]$$

$$\alpha_{j-1}(s_1)$$

$$\gamma_j(s', s_1)$$

$$\beta_j(s_1)$$

$$\beta_{j-1}(s')$$

$$\gamma_j(s', s_2)$$

$$\beta_j(s_2)$$

$$\gamma_j(s', s_2)$$

$$\beta_j(s_2)$$

$$\gamma_j(s', s_2)$$

$$\beta_j(s_2)$$

$$\gamma_j(s', s_2)$$

$$\beta_j(s_2)$$
\[
\beta_{j-1}(s') = \max^* \left[ \beta_j(s) + \gamma_j(s', s) \right] 
\]
(5.2)

\[
\gamma_j(s', s) = \log \left[ p(y_j \mid x[-m, \ldots, j + m])P(x_j) \right] 
\]
(5.3)

The operator max* is defined as \( \max^* = \max(x, y) + \log(1 + e^{-|x-y|}) \). \( P(x_j) \) is the a priori probability of a transmitted bit \( x_j \). Usually an equiprobable distribution of the transmitted bits is assumed. There are two edges coming into and going out of the states in Fig. 5.2.

For the forward recursion the maximum of two possible previous states propagates and for the backward recursion the maximum of two possible states propagates. The forward recursion goes through the states in increasing order in time index \( j=1,2,\ldots,n \) and the backward recursion in decreasing order in time index \( j=n,n-1,\ldots,1 \).

For the proper operation of the BCJR algorithm initialization is required. Let \( s_0 \) be an initial state. Then the forward and backward metrics are initialized as:

\[
\alpha_0(s) = \begin{cases} 
0, & s = s_0 \\
-\infty, & s \neq s_0
\end{cases} \quad \text{and} \quad \beta_0(s) = \begin{cases} 
0, & s = s_0 \\
-\infty, & s \neq s_0
\end{cases} 
\]
(5.4)

### 5.3 Turbo equalizer

The BCJR algorithm calculates the most probable transmitted bit in according to the knowledge of the channel (conditional probabilities) and the received sequence of samples. However the equalizer has the task of determining the reliabilities of the output bits corresponding to the central bits of the terminal states. The bit reliabilities are described by the ratio of conditional probabilities in log domain:
The value of the LLR quantifies the ratio of the probabilities for the received bit being a 0 vs. the received bit being a 1 by dividing the conditional probabilities of the transmitted bit $x_j$ being a 0 or 1 given the received sample $y_j$. If the ratio is greater than 0 the received bit is 1 and if the ratio is smaller than 0 the received bit is 0. Greater absolute value of the LLR indicates greater certainty for the received bit. The conditional probabilities are given by averaging all states in which the middle bit is involved

$$p(y_j \mid x_j = 0,1) = \sum_{s|x_j=0,1} p(y_j \mid s)p(s)$$  \hspace{1cm} (5.6)$$

The calculation the reliabilities is based on the previously described metrics:

$$L(x_j) = \max_{(s',s):x_j=0} \left[ \alpha_{j-1}(s') + \gamma_j(s',s) + \beta_j(s) \right] - \max_{(s',s):x_j=1} \left[ \alpha_{j-1}(s') + \gamma_j(s',s) + \beta_j(s) \right]$$ \hspace{1cm} (5.7)$$

The proposed turbo equalization scheme has a particular architecture which is shown in Fig. 5.3.

![Fig. 5.3 Block diagram of proposed turbo equalization scheme.](image_url)
The BCJR equalizer and the LDPC decoder are connected in an iterative loop. Upon receiving a codeword the equalizer calculates the initial LLRs for the LDPC decoder. After the decoder is finished it calculates extrinsic LLRs and passes them back to the equalizer for the second iteration. The extrinsic LLRs are the difference between two discrete periods of time. They are calculated for every bit of the sequence. The iteration at which the extrinsic reliabilities are passed back to the BCJR equalizer is called outer iteration. The iterations in the LDPC decoder are referred to as inner iterations. The LLR for bit $x_j$ is given by

$$L_{LDPC, ext} (x_j^t) = L_{LDPC} (x_j^t) - L_{LDPC} (x_j^{t-1})$$

(5.8)

where $x_j^t$ indicates is bit $x_j$ at iteration $t$. These extrinsic bit reliabilities are used as a priori probabilities for the BCJR equalizer. The calculation of the a priori reliabilities for subsequent outer iteration is performed by:

$$L_{BCJR, apr} (x_j) = \log \left[ P \left( x_j \right) \right] = L_{LDPC, ext} (x_j)$$

(5.9)

Simulation results for the proposed scheme are discussed next.

### 5.4 Simulations and results

The chromatic dispersion is a linear impairment and can be compensated for in many different ways. A typical method for compensation is the Viterbi equalizer which serves as a hard decision based maximum likelihood sequence detector. The biggest disadvantage of the Viterbi equalizer is that its complexity is exponentially growing and after certain distance the number of states required for proper chromatic dispersion
compensation it becomes too high for practical implementation. For example, compensation for chromatic dispersion accumulated in 1000km of fiber requires 8192 states.

Verification of chromatic dispersion compensation by turbo equalization is performed in the next simulation by observing an NRZ modulation at the rate of 10Gb/s. The dispersion map for the system is presented in Fig. 5.4. (a), while Fig. 5.4 (b) shows the results. Single channel optical system with 700km of SMF28 is used with dispersion of 16ps/nm-km and dispersion slope of 0.08ps/(nm²-km) The NRZ pulses were modeled as raised cosine pulse with roll-off factor of 0.5. The launch power for the simulation was 0dBm. The code LDPC(16935,13550) was studied operating in the proposed turbo equalization scheme with the memory of $2m+1=7$. The figure contains two plots – for the BJCR equalizer and for turbo equalizer. As seen the BCJR equalizer enters an error floor around BER of $5 \times 10^{-3}$. The turbo equalizer is capable of operating without error floor below BER of $5 \times 10^{-9}$.

As mentioned above the fiber nonlinearities are the dominant impairment for optical communication systems operating at 40Gb/s and above. The simulation of nonlinear effects requires the use of realistic model of a fiber optic communication system. The model from [30] has been used for the purpose. This model uses the nonlinear Schrödinger equation for the pulse propagation. It incorporates the Kerr nonlinearities, as well as chromatic dispersion, polarization mode dispersion, stimulated Raman,
scattering, amplified spontaneous emission (ASE) noise, filtering effects, ISI and linear crosstalk effects.

![Dispersion map under study of chromatic dispersion compensation via LDPC-coded turbo equalization](image)

**Fig. 5.4** (a) Dispersion map under study of chromatic dispersion compensation via LDPC-coded turbo equalization (b) Performance of the turbo equalizer for chromatic dispersion compensation for LDPC(16935,13550) code, memory $m+1=7$ and 700km of SMF28.
The dispersion map shown on Fig. 5.5 was designed such that the dominant impairment is the intra-channel four wave mixing. Two types of fiber were used for the simulation \( D_+ \) and \( D_- \). The sign indicates the sign of the dispersion for the fiber of choice. A pre-compensation of \(-1600\) ps was used after the transmitter followed by an optical amplifier then \( N \) sections of 120 km each for long haul transmission.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{dispersion_map.png}
\caption{Dispersion map for dominant intra-channel four wave mixing.}
\end{figure}

The fiber \( D_+ \) has attenuation constant \( 0.19 \text{dB/km} \) with \( 20 \text{ps/nm.km} \) dispersion, dispersion slope of \( 0.06 \text{ ps/(nm}^2\text{km)} \) and nonlinear refractive index of \( 2.6 \times 10^{-20} \text{ m}^2/\text{W} \). The D. fiber has attenuation constant \( 0.2 \text{dB/km} \) with \(-40 \text{ps/nm.km} \) dispersion, dispersion slope of \( 0.06 \text{ ps/(nm}^2\text{km)} \) and the same nonlinear refractive index. Every span is of length 120km consisting of 80 km of \( D_+ \) and 40 km of \( D_- \). The simulations were carried out for RZ modulation format with duty cycle of 33%. Launch power was maintained at 0dBm. The values for the optical and electrical filter were chosen as \( 3R_l \) and \( 0.7R_l \) respectively, where \( R_l \) is the actual transmission rate calculated after dividing the desired data rate by the code rate such that the effective data rate is 40Gb/s.
Fig. 5. 6 BER performance of turbo equalization for different coding schemes with the dispersion map from Fig. 28 Performance of large girth codes with no equalizer vs BCH(128,113)x BCH(256,239).

The simulation results for nonlinearity compensation are shown in Fig. 5.6. The number of spans was used as a parameter. Spans from 20 to 60 in step of 2 were used. The plots are presented as input BER vs output BER. The plot in Fig. 6 shows the performance of the LDPC codes without the turbo equalizer. It is clear from the plot that LDPC codes outperform the standardized code BCH(128,113)x BCH(256,239) for the
same rates. Also it can be seen that LDPC codes of higher girth perform better compared to LDPC codes of lower girth. The block-circulant design LDPC codes outperform the balanced incomplete block design LDPC codes for the same code rate and girth.

**Fig. 5.7** BER performance for different coding schemes for the dispersion map on Fig. 5.5. Large girth LDPC codes vs turbo-product codes and LDPC with equalization.

The results for the same codes with turbo equalizer are shown in Fig. 5.7. The advantage of the turbo equalization scheme is clear after comparing the LDPC vs LDPC in turbo equalization configuration. The same codes perform better in Fig. 5.7 vs Fig. 5.6. Also the effect of the memory is clearly distinguishable. Higher memory assumption results in better turbo-equalization scheme performance. The Large girth LDPC codes
clearly outperform the turbo-product code of the same rate. The larger girth LDPC codes perform better compared to the lower girth LDPC codes.

5.5 Experimental setup and results

To experimentally demonstrate the performance of the proposed equalization scheme several LDPC codes have been tested on the setup shown in Fig. 5.8. A personal computer via GPIB interface controls the pattern generator (Anritsu MP1763C) for pattern upload and the oscilloscope (Agilent 11982A) for data collection.

**Fig. 5.8** Experimental setup.

A zero-chirp Mach-Zehnder modulator (JDSU OC-192) is used for optical modulation of 10Gb/s NRZ electrical signal. Controlled amount of first-order PMD and amplified spontaneous emission (ASE) noise are introduced with the aid of a PMD emulator (JDSU PE3), and an ASE source. Amplification is done with an erbium doped amplifier.
(Optigain, Inc. 2000 series), and photo-detection is performed with a detector Agilent 11982A. Constant level of 0dBm is maintained at the laser source output and the optical signal to noise ratio (OSNR) is controlled by a variable optical attenuator (VOA) and monitored with an optical spectrum analyzer. Constant level of -6dBm is maintained at the detector with the second VOA. The oscilloscope is triggered by the pattern generator. The personal computer serves as an LDPC decoder (the hardware is currently in development stage).

The performance of several large-girth quasi-cyclic LDPC codes was experimentally evaluated. The LDPC codes under study are listed in Table 2.

<table>
<thead>
<tr>
<th>LDPC code</th>
<th>Code length</th>
<th>Code rate, R</th>
<th>The girth, g</th>
<th>Column weight, w</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDPC1</td>
<td>8020</td>
<td>0.950</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>LDPC2</td>
<td>11936</td>
<td>0.906</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>LDPC3</td>
<td>15328</td>
<td>0.906</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>LDPC4</td>
<td>16935</td>
<td>0.800</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>LDPC5</td>
<td>4376</td>
<td>0.936</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>GLDPC</td>
<td>6048</td>
<td>0.81</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

For comparison purposes the random LDPC5 code (MacKay’s random code) is evaluated as well. To guarantee reliable calculation sequence of length approximately 64000 bits is used and acquired multiple times resulting in minimum 10 million test bits per measurement. For every calculation 5 outer and 25 inner iterations were used. The
PMD emulator used is of first order and has limited capabilities up to 125 ps of DGD. This corresponds to 1.25 bit periods. Memory of $m=2$ bits was used for the calculations.

**Fig. 5.9** Performance evaluation of LDPC codes in the presence of constant amount of DGD (125ps).

Figure 5.9 shows the comparison of the BER performance of the aforementioned codes in the presence of DGD of 125 ps for different OSNR values. Quasi-cyclic (QC) LDPC code of girth 6 performs comparable to random MacKay code down to BER of $10^{-4}$ and outperforms it at lower BER values. QC LDPC2 code of $R=0.906$ outperforms the random code by 0.75 dB at BER of $10^{-6}$. QC LDPC3 code of $R=0.906$ outperforms LDPC2 code by 1 dB at BER=$10^{-6}$. QC LDPC4 code of $R=0.8$ outperforms QC LDPC3 code of $R=0.906$ by 2 dB at BER=$10^{-6}$. LDPC4 code of $R=0.81$, $g=10$ outperforms GLDPC code of $R=0.81$, $g=8$ by 0.25 dB at BER=$10^{-6}$. GLDPC code with Hamming
local code performs similarly to code LDPC4 down to BER=10^{-3}. Code LDPC4 outperforms it for lower values of BER. The coding gain for code GLDPC is 8.75dB at BER=10^{-6}.

![Graph showing BER performance for LDPC2 code of rate R = 0.906 for different values of DGD.](image)

**Fig. 5.10** Performance evaluation of LDPC(11936,10819) code of rate $R = 0.906$, for different values of DGD.

Figure 5.10 shows the comparison of the BER performance of LDPC2 code of rate $R=0.906$ for DGD of 0ps, 50ps and 125ps. The OSNR penalty for DGD of 125ps is 3dB at BER=10^{-6}. Coding gain improvement over BCJR equalizer (with memory $2m+1=5$) for DGD=125ps is 6.25dB at BER=10^{-6}. Larger coding gains are expected at lower BERs. Polynomial fit of 4th order was used to obtain the smoothed version of the measurement curve. Notice that girth-10 LDPC codes do not exhibit error floor phenomena down to 10^{-15} [2]. That allows extrapolation of results down to BER of 10^{-15}. Such an extrapolation indicates that the net effective coding gain improvement over
BCJR equalizer for LDPC code is about 12.5 dB at BER of 10^{-15}.

**5.6 Conclusions**

In conclusion, a LDPC-coded turbo equalization was proposed and tested. It operates with soft decision iterative decoding and consists of two major parts – a BCJR equalizer and a LDPC decoder. The equalizer lowers the BER to the error correction threshold of the code and provides log-likelihood reliabilities for the received samples. The LDPC decoder iterates with the equalizer to provide better estimates for the received samples.

For that scheme special class, high rate, girth-10, quasi-cyclic LDPC codes were designed. The codes were tested in a simulated environment for chromatic dispersion and fiber nonlinearities, and experimentally tested in laboratory environment with a physical optical communication system. They were compared to a random MacKay LDPC code, and evaluated for use in PMD compensation by turbo equalization. Properly designed quasi-cyclic code of large girth and sufficient length can match and outperform the random codes. Simulation results indicate that they also outperform the currently standard concatenated BCH codes for chromatic dispersion compensation and the turbo-product codes in the presence of intra-channel four wave mixing. Low decoding complexity implementation is possible due to the low column weight of these codes (3 or 4) and because of low complexity of min-sum-with-correction-term algorithm [72]. Moreover, given the quasi-cyclic structure of the parity-check matrices, the encoder can be implemented using the modulo-2 adders and shift registers. High speed implementation of iterative decoder for this class of LDPC codes can be performed by using either FPGA.
or VLSI. A recent paper [73] describes the first step needed for FPGA implementation of turbo equalizer, namely the FPGA implementation of decoders for large girth LDPC codes.

A significant improvement over the results in [30] has been made with the LDPC code of length 16935, rate $R=0.8$ and girth 10. It provides the net effective coding gain improvement over BCJR equalizer of 9dB at BER=$10^{-6}$ in the presence of 125ps of DGD. All tested codes were within 1.25dB away from capacity limit. GLDPC code of length 6048 and girth 8 performs only 0.25dB worse at BER=$10^{-6}$ compared to LDPC code of length 16935 and girth 10. The rates of the two codes are the same. This is a significant result showing the excellent performance of the GLDPC codes for the used Hamming local code. Better results are expected for local codes of higher strength compared to the Hamming code.

After an extrapolation the net effective coding gain improvement over BCJR equalizer for LDPC4 code is about 12.5 dB at BER of $10^{-15}$. Previous simulation results [65] indicate that no significant improvement in performance of LDPC-coded turbo equalizer is noticed for the number of inner iterations above 25, and the number of outer iterations above 5. The latency of detection/decoding in turbo equalizer can be reduced by reducing the number of outer iterations, without entering the error floor, at the expense of BER performance degradation.
CHAPTER VI - MULTILEVEL BCJR EQUALIZER

6.1 Introduction

Rapid increase of bandwidth demand in the telecommunication industry in recent years has lead to quick evolution of optical communication systems and some areas of research related to them. This trend is most noticeably witnessed by the accelerated growth in the demands for transmission capacity originating from wide variety of services provided over the internet, like audio and video streaming, etc. Network operators are already deploying systems at 40 Gb/s channel transmission per wavelength and upgrading the existing networks to 40 Gb/s. They are also actively researching systems operating at 100 Gb/s per wavelength channel transmission. However there are several major transmission impairments that significantly deteriorate the performance of fiber-optic communication systems at these data rates. The inter-channel and intra-channel nonlinearities, the nonlinear phase noise and the polarization mode dispersion (PMD) \[74]-[77] represent the current limiting factors in the efforts to increase the capacity/speed, extend the transmission distance, and provide more flexible wavelength switching and routing capabilities in optical networks. Novel advanced techniques and devices in modulation and detection, coding and signal processing need to be developed to address the above challenges.

A popular way to increase capacity is increasing the spectral efficiency of transmission by employing multilevel modulation formats. Currently this is a research
topic of high interest and importance. Systems using multilevel modulation schemes operate at much lower transmission symbol rates, which decreases the operating speed for all related signal processing, detection and decoding. It also facilitates the management of PMD compensation and the mitigation of fiber nonlinearities. Most multilevel systems use coherent detection because it retains the phase information of the optical signal.

To deal with PMD and chromatic dispersion a number of equalization techniques have been proposed in the recent past [78],[79]. Some of them are based on digital filtering techniques, while others use maximum likelihood sequence detection (MLSD) or turbo-equalization. Simultaneous suppression of chromatic dispersion and PMD is possible with orthogonal frequency division multiplexing OFDM [80],[81] or turbo-equalization. However OFDM systems do not perform well in the presence of nonlinearities. It has been previously shown in [47],[69],[70] that low-density parity-check (LDPC)-coded turbo equalization is able to simultaneously compensate for multiple impairments, including chromatic dispersion, PMD, noise and fiber nonlinearities in binary fiber-optic transmission systems. Despite being an excellent candidate for simultaneous impairment mitigation this approach is applicable only to binary transmission and direct detection and multilevel systems require novel equalizers that can also operate with coherent detection.

The concept of LDPC-coded turbo-equalization has been generalized in this chapter to multilevel coded-modulation schemes. The proposed multilevel turbo equalizer is
universal and applicable to any two-dimensional signal constellation such as $M$-ary phase-shift keying (PSK), $M$-ary quadrature amplitude modulation (QAM) or $M$-ary pulse-amplitude modulation (PAM), and both coherent and direct detections.

We propose the multilevel maximum a posteriori probability (MAP) turbo equalization scheme based on multilevel Bahl-Cocke-Jelinek-Raviv (BCJR) [82] algorithm (called here multilevel BCJR equalizer) and an LDPC decoder. Again the proposed modulation scheme is suitable for simultaneous compensation of multiple impairments like chromatic dispersion, PMD and nonlinearities as well as imbalance in polarization, phase or power among the different channels. The multilevel BJCR equalizer serves as nonlinear intersymbol interference (ISI) equalizer, and provides the soft symbol log-likelihood ratios (LLRs), which are used in the decoding process. The turbo-equalization process is an iterative exchange of soft information between the LDPC decoder and the equalizer which results in improved decoding performance. Compared to the binary case here the calculation of symbol LLRs is required. The symbol LLRs are then used to calculate the bit LLRs and they in turn are used for the LDPC decoding. Another unique property of proposed scheme is that it considers the independent symbols transmitted in horizontal- (H-) and vertical- (V-) polarizations as a super-symbol. The two polarizations are not considered as independent of each other and the crosstalk energy between the channels is taken into account in the decoding process to improve the overall equalizer performance.

To experimentally verify that concept two independent phase-shift keying (BPSK)
signals at 10Gb/s were used to drive two optical modulators operating in orthogonal polarizations. The created signals were labeled as H- and V-polarization channels. After combining them, controlled amount of PMD and amplified spontaneous noise were added followed by transmission. NRZ modulation format was used for comparison purposes to the performance of the binary BJCR case of the turbo equalizer in a NRZ transmission system with direct detection. We show that in the presence of first order PMD with differential group delay (DGD) of 100ps the OSNR penalty, at bit error ratio (BER) of $10^{-6}$, is less than 2.5 dB, which is smaller by 0.5dB compared to the PMD penalty found in an NRZ transmission system at 10Gb/s [69].

The increase of the spectral efficiency in optical communication systems is important for facilitating the upcoming needs of the telecommunication industry. Some ways to achieve this goal involve multilevel modulation formats and non-binary codes. As an important part of any receiver the proposed multilevel turbo equalizer will be useful for both. As mentioned above it is applicable to any two-dimensional signal constellation for both coherent and direct detections. It can also be utilized in the decoding of non-binary LDPC codes in a turbo equalization fashion similar to the described in [69].

### 6.2 Polarization multiplexed multilevel LDPC-coded modulation and multilevel turbo equalizer

Fig. 6.1 shows the proposed polarization multiplexed multilevel LDPC-coded modulation scheme. On the transmitter side a multilevel signal is generated. In the current experiment $M$-ary PSK was used but the signal can be either $M$-ary QAM or $M$-
ary PAM as well. Every symbol in the multilevel signal carries \( \log_2 M = b \) bits. To create the multilevel signal, \( b \) source bit streams are used. Every bit stream is encoded separately and independently of the rest. The used code is the LDPC code of rate \( R = k/n \), where \( k \) is the information word length and \( n \) is the codeword length. At every symbol period time instance the block-interleaver creates a vector of \( b \) bits and passes it to the mapper.

![Diagram](image)

**Fig. 6.1** Polarization multiplexed multilevel LDPC-coded modulation: (a) transmitter architecture, and (b) receiver architecture. CWL: continuous wave laser, PBS: polarization beam splitter, PBC: polarization beam combiner.

The mapper accepts the vector column-wise and determines the constellation points for transmission by using the appropriate mapping rule (natural, Gray, etc). In the current experiment, the outputs of the mapper are used to drive the modulating signals of the phase modulators (PM). In \( M \)-ary PSK, the outputs of the mapper are used as the modulating signals of the phase modulators (PMs), with \( \phi \in [0, 2\pi/M, 4\pi/M, \ldots, (M-1)\pi/M] \).
1) $2\pi/M$, $l \in \{H,V\}$ (H denotes the horizontal polarization, and V denotes the vertical polarization). The phase modulated signals are then combined and transmitted in orthogonal polarizations over the channel. On the receiver side the signal is split into two signals in orthogonal polarizations – H and V. A local laser signal split into V and H polarizations is used to create signal for coherent detection. Next joint MAP detection with soft iterative decoding is performed. The multilevel BCJR equalizer serves as a MAP equalizer and provides soft log likelihood reliabilities for the LDPC decoder. The LDPC decoder provides extrinsic information to the BCJR decoder for optimized performance. As mentioned before the multilevel BJCR equalizer considers both polarizations as a super-symbol. It calculates symbol LLRs instead of bit LLRs. This is why an additional block for bit LLR calculation is required before the LDPC decoder block. The bit LLR calculation block provides $b$ bit streams in a column vector which are fed into $b$ binary LDPC decoders. The output of every decoder provides information for the calculation of the extrinsic symbol LLRs fed back into the multilevel BCJR equalizer.

In order to match the LDPC decoders to the BCJR equalizer we use the approach of extrinsic information transfer (EXIT) charts described in [71].

The multilevel BCJR equalizer is described in more detail next. Similar to the binary BCJR equalizer, it operates on an optical channel model described by discrete dynamical trellis. The used channel model is a nonlinear ISI channel with memory $2m+1$. The trellis operates on the received super-symbol sequence in order to determine the most probable transmitted super-symbol sequence. The transmitted super-symbol sequence is denoted by $X$ and the received super-symbol sequence is denoted by $Y$. The super-symbol
$\mathbf{x} \in \mathbf{X}$ consists of two components $\mathbf{x} = (s_H, s_V)$. The first one $s_H$ denotes the symbol transmitted over the horizontal polarization and the second one $s_V$ denotes the symbol transmitted over vertical polarization. The received symbol $\mathbf{y} \in \mathbf{Y}$ has four components $\mathbf{y} = (\Re(y_H), \Im(y_H), \Re(y_V), \Im(y_V))$. $\Re(y_H)$ and $\Im(y_H)$ denote the samples corresponding to I- and Q-channels in horizontal polarization, while $\Re(y_V)$ and $\Im(y_V)$ denote the samples corresponding to I- and Q-channels in vertical polarization. $y_j$ is the vector of samples that correspond to the transmitted symbol $x_j$ for the $j^{th}$ discrete time instance. The memory assumption indicates that a super-symbol $x_i$ is influenced by total of $2m$ super-symbols - the preceding $m$ ($x_{i-m}, x_{i-m+1}, \ldots, x_{i-1}$) super-symbols and the next $m$ ($x_{i+1}, \ldots, x_{i+m}$) super-symbols in the sequence. $m$ is considered to be the memory of the state, where a state is defined as $\mathbf{s} = (x_{i-m}, \ldots, x_i, x_{i+1}, \ldots, x_{i+m})$. The main difference in this channel model compared to other channel models is that both previous and following symbols are considered to influence the current symbol. The set of all possible states at a moment in time $j$ define a vector of states. The set of $2m+1$ vectors for the corresponding $2m+1$ memory super-symbols define a trellis. At any moment of time the triple {previous state, channel output, next state} uniquely define the discrete trellis for a given state. The multilevel BCJR algorithm operates on the discrete trellis at every moment in time.

To visualize that concept an example trellis for the polarization multiplexed PSK modulation used in the experiment is illustrated on Fig. 6.2 (a). It shows the 4-level ($M=4$) trellis for memory $2m+1=3$, at time instant $j$ and at time instant $j-1$, meaning that the left-column denotes the previous instance of time, and the right-column the current instance of time. The quaternary numerical system is used for this example (base 4). The
following instance in time or \( j+1 \) is not represented in the diagram.

![Trellis Diagram](image)

**Fig. 6.2** (a) 4-level trellis example for channel description, at time instant \( j \). (b) Constellation diagram for the used modulation format.

The trellis has \( M^{2m+1} = 64 \) states, each corresponding to the different possible 3-symbol patterns. Let \( S \) denote that set \( S = \{s_0, s_1, s_2, \ldots, s_{63}\} \). The triple of digits next to every state represents the previous, the current and the following super-symbols for that state in the quaternary numerical system. Every symbol can take a value from 0 to 3. For the current experiment two BPSK signals were used with natural mapping. The following sequences of bits 00, 01, 10, 11 were mapped into the constellation points (CP) CP(0)=(0,0),
CP(1)=(1,0), CP(2)=(0,1), CP(3)=(1,1). The constellation diagram is shown in Fig. 6.2 (b). The middle super-symbol of the current state indicates the output super-symbol, which is to be detected. The edge arrows indicate possible transitions from a given state in a discrete moment of time to a state in the next discrete moment of time. There are two labels associated with every transition edge arrow. The first one indicates the currently transmitted super-symbol and the second one indicates the super-symbol in terminal state for the current transition. The transition edges originating from every state are equal to \( M \). The same is valid for the number of transition edges merging in a terminal state. Edges merging in the same state have the same arrow pattern for easier comprehension of the diagram.

To completely characterize the optical channel, and respectively the trellis on which it operates, the conditional density functions
\[
p(y_j|s_j) = p(y_j|s), \quad s \in S
\]
are needed. They can be determined by using the instanton-Edgeworth expansion method [83]. The second method, which is used in the current experiment, calculates histograms after transmitting and collecting samples of a sufficiently long training sequence.

Fig. 6.3 shows a graphical representation of the update rule for the metrics for the example trellis on Fig. 6.2 (a). They are defined in a similar fashion to the metrics of the binary BCJR equalizer [82]. As mentioned above the numbers of edges originating from and merging into every state are equal to the signal constellation size \( M \). The forward metric is defined as the probability of the terminal state being \( s \) given a received sequence \((y_1, y_2, \ldots, y_j)\). In log-domain the corresponding reliability is given by
\[
\alpha_j(s) = \log[p(s_j=s, y_1 \ldots y_j)], \quad (j=1,2 \ldots n).
\]
The backward metric is related to the conditional probability of
receiving the next \( n-j \) symbols given the current state.

\[
\alpha_j(s') \circ \gamma_j(s', s) \quad s' \quad \alpha_j(s) = \max_{s'} \left[ \alpha_{j-1}(s') + \gamma_j(s', s) \cdot \alpha_{j-1}(s'_{2}) + \gamma_j(s'_{2}, s), \quad \alpha_{j-1}(s') + \gamma_j(s', s) \cdot \alpha_{j-1}(s'_{4}) + \gamma_j(s'_{4}, s) \right]
\]

\[
\beta_{j-1}(s') = \max_{s} \left[ \beta_j(s_1) + \gamma_j(s', s_1), \beta_j(s_2) + \gamma_j(s', s_2), \beta_j(s_3) + \gamma_j(s', s_3), \beta_j(s_4) + \gamma_j(s', s_4) \right]
\]

**Fig. 6.3** (a) Forward recursion calculation for 4-level BCJR equalizer (b) Backward recursion calculation for 4-level BCJR equalizer.

In log-domain the reliability is defined by \( \beta_j(s) = \log[p((y_{j-1} \ldots y_n) | s_j = s)] \). The branch metric is defined as the transition probability between two states when a symbol \( y_j \) has been received. The corresponding reliability in log-domain is defined by \( \gamma_j(s', s) = \log[p(s_j = s, y_j, s_{j+1} = s')] \). The metrics have to be updated iteratively for every discrete period of time as given below:
The operator \( \text{max}^* \) is defined as \( \text{max}^*(x,y) = \text{max}(x,y) + \log(1 + e^{-|x-y|}) \). To efficiently calculate the operator one can use \( \text{max}^*(x,y) = \text{max}(x,y) + cf(x,y) \), where \( cf(x,y) = (1 + e^{-|x-y|}) \) is called correction factor. It is usually approximated and implemented as a lookup table.

\( P(x_j) \) is the a priori probability of a transmitted symbol \( x_j \). Usually an equiprobable distribution of the transmitted symbols is assumed. Because \( M=4 \) there are 4 edges coming into and going out of the states in Fig. 6.3. The forward recursion goes through the states in increasing order in time index \( (j=1,2,...,n) \) and the backward recursion in decreasing order in time index \( (j=n,n-1,...,1) \). For every discrete moment of time the metrics are calculated and stored. First pass is with forward time index, when \( \alpha_j \) are calculated and the second pass is with backward time index, when \( \beta_j \) are calculated. The branch metrics are common for both the forward and the backward passes and are calculated only once and stored. As in the binary case of the BJCR algorithm the forward and backward metrics are initialized as:

\[
\alpha_0(s) = \begin{cases} 
0, s = s_0 \\
-\infty, s \neq s_0
\end{cases} \quad \text{and} \quad \beta_n(s) = \begin{cases} 
0, s = s_0 \\
-\infty, s \neq s_0
\end{cases}
\]  

where \( s_0 \) is an initial state.

In the multilevel case, we determine symbol reliabilities instead of bit reliabilities.
They are used to calculate the reliabilities for the bits within the symbol. Bit reliabilities are then used in subsequent LDPC decoding. To estimate the symbol LLRs we use the following relationship:

\[
\Lambda(x_j = \delta) = \max_{(s',s) : x_j = \delta} \left[ \alpha_{j-1}(s') + \gamma_j(s',s) + \beta_j(s) \right] - \max_{(s',s) : x_j = \delta^0} \left[ \alpha_{j-1}(s') + \gamma_j(s',s) + \beta_j(s) \right] \quad (6.5)
\]

In equation (3) the symbol \( \delta \) is the candidate output symbol, and the reliability indicates the likelihood of the received symbol being a particular transmitted symbol \( \delta \in \{x_1, x_2, \ldots, x_{M-1}\} \). In a binary modulation format a scalar reliability represents every bit, while multilevel modulation formats require a vector of reliabilities of symbols. The use of a reference symbol \( \delta_0 = x_0 \) reduces the size of the reliability vector from \( M-1 \) to \( M \). Once the symbol reliabilities have been calculated, bit reliabilities can be determined in log domain by observing the binary representation of every symbol. Let \( c_k \) represents the \( k^{th} \) bit in the binary representation of a symbol \( x_j \), \( k=1,2,\ldots,b \). Then the likelihoods for the bit \( c_k \) in log domain can be calculated using the already calculated likelihood for all symbols \( x_j \) as follows:

\[
L(\hat{c}_k) = \log \left[ \sum_{x_j, c_k = 0} e^{(\Lambda(x_j))} \right] - \log \left[ \sum_{x_j, c_k = 1} e^{(\Lambda(x_j))} \right] \quad (6.6)
\]

The summation on the left is performed over all symbols \( x_j \) that have 0 in \( k^{th} \) position and the summation on the right is performed over all symbols \( x_j \) that have 1 in \( k^{th} \) position. Once the bit reliabilities have been calculated the inner iterations of the LDPC decoder are performed based on the min-sum-with-correction-term algorithm.

Now what is left is to calculate the extrinsic bit reliabilities used for improvement.
of the performance of the LDPC-coded turbo equalizer. Similarly to the binary case of
the algorithm extrinsic reliabilities are calculated as the difference between the bit
reliabilities of the current $t$ and the previous iteration $t-1$ as given below

$$L_{LDPC,ext}(c^t_k) = L_{LDPC}(c^t_k) - L_{LDPC}(c^{t-1}_k)$$  \hspace{1cm} (6.7)

where $c^t_k$ indicates is the $k^{th}$ bit of symbol $x_j$ at iteration $t$. These extrinsic bit reliabilities
are used to calculate the extrinsic symbol reliabilities, which are used as a priori
probabilities for the multilevel BCJR equalizer. The calculation of the a priori symbol
reliabilities for subsequent outer iteration is performed by:

$$L_{BCJR,apr}(x_j) = \sum_{k=0}^{l-1} (1-c_k) L_{LDPC,apr}(c_k) = \log[P(x_j)]$$  \hspace{1cm} (6.8)

6.3 Experimental setup and results

Fig. 6.4 shows the experimental setup being used for the verification. The dotted
lines indicate the offline processing for the proposed turbo-equalization scheme. A
continuous wave laser signal is passed through a polarization beam splitter resulting in
two new continuous wave laser signals in orthogonal polarizations. Each beam is then
modulated by a phase modulator (Covega). The used modulator driver amplifiers were
JDSU H301-1110. Two independent driver signals are used to drive the modulators. They
are generated by a pulse generator at 10 Gb/s (Anritsu MP1763C). The used sequences
were LDPC-coded in a PC. They have equal number of 0s and 1s. The loading of the
sequences into the pattern generator was done via personal computer with GPIB
interface.
Fig. 6.4 Experimental setup for polarization multiplexed BPSK. CW Laser: continuous wave laser, PM: phase modulator, ASE: amplified spontaneous emission noise source, 3dB: 3dB coupler.

The two phase-modulated signals are combined by using a polarization beam combiner. Instead of polarization maintaining fiber, regular SMF28 fiber and polarization controllers were used to maintain the polarization of the signal. Polarization controllers are not shown on the figure due to limited space. PMD was used as signal impairment for this experiment. A JDSU PE3 emulator was used to introduce controlled amount of first order PMD. The noise was generated by amplified spontaneous emission (ASE) source added to the signal with a 3dB coupler. A variable optical attenuator (VOA) was used to control the level of the noise. The level of the modulated signal was maintained at 0dBm while the ASE power level was varied to obtain the required optical signal-to-noise ratios (OSNRs). The measurement of the OSNRs was performed with optical spectrum analyzer (OSA) (Agilent 86142B) with 0.1nm reference bandwidth. Next the signal is amplified with optical amplifier (Optigain, Inc. 2000 series). The amplifier is not shown due to
limited space. On the receiver side the signal is first filtered (JSDU 2nm band-pass filter) and then processed through a polarization beam splitter. Polarization controllers are used to maintain the polarizations between transmitter and receiver. To perform coherent detection the local oscillator signal was split with polarization beam splitter and mixed with the received signal by 3dB coupler. The resulting signals are detected with two detectors (Agilent 11982A) and (Agilent 86105A). The signals from the detectors are sampled by digital oscilloscope (Agilent DCA 86100A). The oscilloscope is triggered by the data pattern via synchronization with the pulse generator. To maintain constant power of -6 dBm at the detector, a VOA was used. Data was transferred via the GPIB interface back to the PC. The PC also served as an offline multilevel turbo equalizer and LDPC decoder.

For the experiment, a quasi-cyclic LDPC(16935, 13550) code of girth 10 and column weight 3, was used. The number of extrinsic iterations between LDPC decoder and BCJR equalizer was set to 3, and the number of the intrinsic LDPC decoder iterations was set to 25. Memory of $2m+1=3$ was sufficient for the compensation of the first order PMD with DGD of 100 ps. Due to extensive memory requirements for larger constellations and larger DGD values a reduced version of the multilevel BCJR algorithm would be required. For example, each codeword can be split into smaller sub-sequences that are processed independently of each other. Also not all branch metrics have to be memorized but only the ones that are largest thus resulting in modified updating rule that takes into account only the dominant branches in the trellis.
Fig. 6.5 BER performance of proposed polarization multiplexed LDPC-coded modulation scheme.

The experimental results for joint performance of the multilevel LDPC-coded turbo equalizer are summarized in Fig. 6.5. The BER penalty due to PMD with DGD of 100ps is 1.5 dB at BER of $10^{-6}$. Coding gain for DGD of 0 ps is 7.5dB at BER of $10^{-6}$, and the coding gain for DGD of 100ps is 8 dB. The polarization separation between the channels was measured to be 20 dB for DGD of 100 ps, and 25 dB for DGD of 0 ps. All the results are reported for OSNR per bit which is 3dB lower than the OSNR per symbol.

The BER performance, when polarization channels were considered independently, is shown in Fig. 6.6 (a) and Fig. 6.6 (b) for DGD of 0 ps and DGD of 100 ps respectively.
Fig. 6.6 BER performance of separate polarization channels utilizing LDPC-coded turbo equalization: (a) DGD=0 ps, and (b) DGD=100 ps.

The coding gain compared to the joint decoding case is 5.5dB at BER $10^{-6}$ with respect to
the uncoded curve for the jointly decoded case. The uncoded curves for DGD of 100ps exhibit a BER floor. We see that the joint detection-decoding approach significantly outperforms the separate decoding case, by even 5dB at BER of $10^{-6}$ for DGD of 100ps. Therefore, the joint equalization and decoding is also successful in compensation for any unbalance in the signal distribution between channels. We observe that the penalty due to DGD of 100ps for separate polarization channels is about 1dB.

For comparison purposes the results for the same code with direct detection, NRZ amplitude modulation and 10Gb/s data rate are used. The multilevel modulation scheme with coherent detection outperforms the binary modulation scheme with direct detection by more 2.75dB at BER of $10^{-6}$. The DGD of 100ps was used for the multilevel case while DGD of 125ps was used for the binary system. Also the memory of the equalizer for the multilevel case was 1 while memory of 2 was used for the binary case of the LDPC-coded turbo equalization. This is very significant improvement in performance over the binary system.

We observe that the binary system with direct detection outperforms the independently considered channels from the multilevel system and coherent detection by more than 3dB at BER $10^{-6}$. This again indicates that when polarization separated signals are treated as independent of each other the two signals are significantly deteriorated.

### 6.4 Conclusions

In conclusion, we proposed a particular polarization multiplexed coded modulation scheme, which considers independent symbols transmitted over both polarizations as a
super-symbol. This scheme is based on coherent detection and employs multilevel BCJR equalizer and LDPC codes as channel codes. For 20 Gb/s transmission, the penalty (at BER of $10^{-6}$) due to PMD and 100 ps of DGD is found to be less than 1.5 dB, while the coding gain is 9dB.

We have experimentally demonstrated that polarization separated channels have to be considered as dependent. We have also shown that the proposed equalization can significantly improve the performance of the system by compensating for that dependence as well as for any imbalance in polarization and phase. The improvement for DGD of 100ps at BER of $10^{-6}$ for joint equalization-decoding over separate equalization-decoding is 5dB. The proposed approach is universal and can be applied to different multilevel coded-modulation schemes.
CHAPTER VII - CONCLUDING REMARKS AND FUTURE WORK

In the current work, we present several methods for improving system performance of high-speed fiber-optics communication systems in the presence of linear and nonlinear signal impairments. The full advantage of LDPC-coded modulation schemes can be exploited only when reliable soft log-likelihood ratios (LLRs) of bits are available. These reliabilities are obtained by employing the maximum \textit{a posteriori} probability Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm for equalization. Significant improvements in system performance are reported. Simulations and experimental results confirm the ability of this scheme to successfully mitigate the effects of multiple impairments simultaneously. The new turbo-equalization scheme is used as a means to simultaneously mitigate both linear and nonlinear channel impairments. This approach is general and applicable to both direct and coherent detection. To further improve the tolerance to different linear and nonlinear channel impairments, we perform the iteration of extrinsic LLRs between BCJR equalizer and LDPC decoder.

In this dissertation, we also provide a comprehensive study of LDPC codes that are suitable for use in high-speed optical communication systems. Different code designs are discussed and their bit error rate (BER) performance versus optical signal to noise ratio (OSNR) is investigated. The results are compared among each other and certain types of codes are selected as promising candidates for future hardware implementation.
For the purpose of optimizing system performance a quantitative measure of the channel capacity is a necessity. In this work a method, based on the forward step of the BCJR algorithm, is presented and employed in practically measuring the achievable information rates (AIRs) of an optical channel. The AIR is the lower bound of the channel capacity. We provide simulations and experimental results for the effects of different impairments on the channel capacity as well as study the influence of quantization effects. We demonstrate experimentally that by using LDPC codes near maximum transmission channel capacity is achievable.

An effective way to increase transmission bandwidth is to increase the spectral efficiency of the signal, by employing the multilevel modulation schemes. In multilevel modulation schemes, multiple electrical digital streams are encoded into the optical pulses and for the same baud rate, the aggregate information rate can be increased. The same binary transmission rate becomes achievable for lower symbol rate and all related signal processing is performed at the lower symbol rate. At lower symbol rates it is easier to compensate for PMD and to mitigate fiber nonlinearities. We propose a novel multilevel *a posteriori* probability turbo equalization scheme based on multilevel BCJR algorithm and LDPC decoder. This turbo equalizer considers independent symbols transmitted over orthogonal polarizations as two-dimensional super-symbols. We show that the proposed multilevel scheme significantly outperforms the corresponding turbo equalizer that considers symbols transmitted in different polarizations as independent.

Future work includes the experimental validation of the proposed turbo equalization
scheme and multilevel turbo equalization scheme in the presence of chromatic dispersion, fiber nonlinearities and in the presence of different combinations of simultaneous impairments. The investigation of multilevel modulation formats for increased spectral efficiency transmission is another area that can benefit from this study. The investigation of systems with very high spectral efficiency, for example 10 bits/s/Hz, in the presence of nonlinearities, chromatic dispersion and PMD, as well as combinations of them and the evaluation of performance of the proposed multilevel turbo equalization in such schemes are of high practical importance.
REFERENCES


[83] M. Ivkovic, I. B. Djordjevic, P. Rajkovic, and B. Vasic, "Pulse energy probability density functions for long-haul optical fiber transmission systems by using