

University College London
Department of Electronic and Electrical Engineering

Digital Signal Processing for Coherent Optical Fibre Communications

David Samuel Millar

A thesis submitted for the degree of Ph.D.

Supervisor: Dr S.J. Savory

I, David Samuel Millar confirm that the work presented in this thesis is my own.
Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

In this thesis investigations were performed into digital signal processing (DSP) algorithms for coherent optical fibre transmission systems, which provide improved performance with respect to conventional systems and algorithms. Firstly, an overview of coherent detection and coherent transmission systems is given.

Experimental investigations were then performed into the performance of digital backpropagation for mitigating fibre nonlinearities in a dual-polarization quadrature phase shift keying (DP-QPSK) system over 7780 km and a dual-polarization 16-level quadrature amplitude modulation (DP-QAM16) system over 1600 km. It is noted that significant improvements in performance may be achieved for a nonlinear step-size greater than one span. An approximately exponential relationship was found between performance improvement in Q-factor and the number for required complex multipliers.

DSP algorithms for polarization-switched quadrature phase shift keying (PS-QPSK) are then investigated. A novel two-part equalisation algorithm is proposed which provides singularity-free convergence and blind equalisation of PS-QPSK. This algorithm is characterised and its application to wavelength division multiplexed (WDM) transmission systems is discussed.

The thesis concludes with an experimental comparison between a PS-QPSK transmission system and a conventional DP-QPSK system. For a 42.9 Gb/s WDM system, the use of PS-QPSK enabled an increase of reach of more than 30%. The resultant reach of 13,640 km was, at the time of publication, the longest transmission distance reported for 40 Gb/s transmission over an uncompensated link with standard fibre and optical amplification.

Acknowledgements

First and foremost I would like to thank my primary and secondary supervisors, Dr. Seb Savory and Prof. Polina Bayvel for the guidance, assistance and wisdom that they have provided me during my studies at UCL. I would also like to thank my colleagues, collaborators and friends in the Optical Networks Group and elsewhere. Particularly valuable in their contributions and discussions have been: Carsten Behrens who assisted with simulations and experimental work; Dr. Yannis Benlachtar who provided stimulating discussions on OFDM systems and nonlinear transmission; Dr. Steve Desbruslais who was a source of expertise on physics, submarine systems and information theory; Dr. Giancarlo Gavioli who built the recirculating fibre loop used in the experiments in this thesis; Dr. Robert Killey who provided insight and feedback into the theoretical aspects of this work; Domaniç Lavery who assisted with both optimising software and performing experiments; Dr. Sergejs Makovejs who designed the QAM16 transmitter used in this work and assisted with all experiments; Jose Mendinueta who provided discussion and debate on mathematical and software issues; and Dr. Benn Thomsen who provided discussion and ideas on theoretical and experimental matters.

List of Abbreviations

ADC	Analogue to digital convertor
ASE	Amplified spontaneous emission
ASIC	Application specific integrated circuit
AWGN	Additive white Gaussian noise
AWG	Arrayed waveguide grating
BER	Bit error rate
CD	Chromatic dispersion
CMA	Constant modulus algorithm
CMOS	Complementary metal-oxide-semiconductor
DBP	Digital backpropagation
DD	Decision directed
DFB	Distributed feedback
DGD	Differential group delay
DP	Dual polarization
DQPSK	Differential quadrature phase shift keying
DSP	Digital signal processing
EDFA	Erbium doped fibre amplifier
FEC	Forward error correction
FIR	Finite impulse response
GBd	Giga baud
GbE	Gigabit Ethernet
GVD	Group velocity dispersion
IM-DD	Intensity modulation-direct detection
ISD	Information spectral density
ISI	Intersymbol interference
LMS	Least mean squares
LO	Local oscillator
ML	Maximum likelihood
MSE	Mean square error
MZM	Mach-Zehnder modulator
MZI	Mach-Zehnder interferometer
NRZ	Non-return to zero

OSNR	Optical signal to noise ratio
PDM	Polarization division multiplexed
PMD	Polarization mode dispersion
PS	Polarization-switched
PSK	Phase shift keying
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
RIN	Relative intensity noise
RDE	Radially directed equaliser
RLS	Recursive least squares
Rx	Receiver
RZ	Return-to-zero
SE	Spectral efficiency
SNR	Signal-to-noise ratio
Tx	Transmitter
WDM	Wavelength division multiplexing

Contents

1	Introduction.....	14
1.1	Introduction	14
1.2	Chapter Overview	17
1.3	List of Publications	18
1.4	Main Contributions	20
2	Literature Review and Theory	21
2.1	Abstract	21
2.2	Introduction	22
2.3	Coherent Detection.....	23
2.3.1	Phase and Polarization Diverse Coherent Receiver	24
2.3.2	Local Oscillator Phase and Frequency Locking.....	26
2.3.3	Analogue to Digital Conversion.....	27
2.4	Modulation	27
2.4.1	Modulation Formats and Coding.....	27
2.4.2	Phase Shift Keying	29
2.4.3	Quadrature Amplitude Modulation	29
2.4.4	Bit-to-Symbol Mapping.....	30
2.4.5	Comparison of the Performance of Modulation Formats.....	32
2.4.6	Optical Modulators	34
2.5	Fibre Transmission Impairments.....	37
2.5.1	Chromatic Dispersion.....	37
2.5.2	Polarization Mode Dispersion	37
2.5.3	The Kerr Nonlinearity	38
2.5.4	Additive Noise.....	38
2.5.5	Laser Phase Noise.....	39

2.6	Digital Signal Processing Algorithms for Coherent Detection	39
2.6.1	Resampling	40
2.6.2	Signal Preparation	40
2.6.3	Stationary Inverse Channel.....	40
2.6.4	Adaptive Equalization Algorithms	42
2.6.5	Least Mean Squares Algorithm	44
2.6.6	Frequency Offset Compensation Algorithms.....	46
2.6.7	Carrier Phase Estimation Algorithms.....	48
2.7	Digital Backpropagation	50
2.7.1	Nonlinear Channel Models.....	50
2.7.2	Split-Step Methods	52
2.7.3	Two- and Three- Block Nonlinear Models	53
2.7.4	Channel Models and Accuracy Considerations.....	54
2.7.5	Digital Backpropagation Algorithms	55
2.7.6	Recent Experimental Results in Nonlinear Backpropagation	57
2.7.7	Alternative Approaches to Nonlinearity Compensation.....	58
2.8	Summary	59
3	Experimental Analysis of Digital Backpropagation	61
3.1	Abstract	61
3.2	Introduction	62
3.3	Experimental Transmission Setup.....	64
3.4	Simulation of Nonlinear Equaliser Performance	66
3.5	DP-QPSK Transmission Results.....	67
3.6	DP-QAM16 Transmission Results.....	73
3.7	Comparison of Performance for DP-QPSK and DP-QAM16.....	77
3.8	Estimated Implementation Complexity for Nonlinear Backpropagation.....	80
3.9	Summary	82

4	Polarization Switched QPSK: Theory and Digital Equalisation.....	84
4.1	Abstract	84
4.2	Introduction	85
4.3	PS-QPSK Modulation and Coding Performance	87
4.4	The Polarization Switched CMA Equaliser	93
4.5	Comparison of the PS-CMA with Generic Equalisers.....	95
4.6	Characterisation of PS-CMA Performance.....	99
4.7	Comparison with Other Published PS-QPSK Algorithms	103
4.8	Application of PS-QPSK to 100 GbE WDM Systems.....	105
4.9	Summary	108
5	Generation and Long-Haul Transmission of Polarization-Switched QPSK at 42.9 Gb/s	109
5.1	Abstract	109
5.2	Introduction	110
5.3	PS-QPSK Generation and Experimental Setup.....	111
5.4	Digital Signal Processing for Experimentally Generated PS-QPSK Signals	113
5.5	Back-to-Back and WDM Transmission Results at 42.9 Gb/s.....	115
5.6	Summary	117
6	Conclusions and Topics for Further Research.....	118
6.1	Abstract	118
6.1.1	Boundaries on Performance Benefits resulting from Nonlinearity Compensation	118
6.1.2	DSP Algorithms for Optimal 4D Modulation Formats	118
6.1.3	Equaliser Initialisation Algorithms.....	119
6.2	Conclusions	119
7	References.....	123

Figures

Figure 2.1 - Quadrature coherent detection.....	23
Figure 2.2 - Phase and polarization diverse coherent receiver. Dashed lines show signals required only when balanced detection is used.	25
Figure 2.3 - Constellation diagrams for BPSK (left), QPSK (centre) and 8PSK (right). Noise loaded to E_s/N_0 of 24.5 dB.	29
Figure 2.4 - Square QAM16 (left), and QAM64 (right). Noise loaded to E_s/N_0 of 24.5 dB.....	30
Figure 2.5 - Star QAM8 (left), and offset star QAM8 (right). Noise loaded to E_s/N_0 of 24.5 dB.	30
Figure 2.6 - QPSK constellations with standard binary coding (left), and Gray coding (right).	31
Figure 2.7 - Noise limited receiver performance for various modulation formats. ...	33
Figure 2.8 - Push-pull MZM operation.	35
Figure 2.9 - MZM transfer function.	35
Figure 2.10 - Dual-polarization triple Mach-Zehnder modulator.	36
Figure 2.11 - DSP Signal Flow Model.....	40
Figure 2.12 - Block diagram of a Bussgang equaliser.	43
Figure 2.13 - The 2x2 MIMO Bussgang Equaliser, figure taken from [46].	43
Figure 2.14 - The Wiener model.....	54
Figure 2.15 - The Hammerstein model.....	54
Figure 2.16 - The Wiener-Hammerstein model.....	54
Figure 3.1 - Recirculating loop setup used for transmission experiments, with optical front end of the phase and polarization diverse digital coherent receiver.	64
Figure 3.2 - Transmitter structure for DP-QPSK, with optional QAM16 stage highlighted. Inset: optical eye-diagrams at the output of the transmitter for DP-QPSK (top), and DP-QAM16 (bottom).	65
Figure 3.3 - Contour plot of experimentally determined Q-factor in dB against launch power and nonlinear step-size for Wiener cascade compensation of 97 spans transmission DP-QPSK at 10.7 GBd. Nonlinear step-size of a single span lies at 80.2 km.....	68
Figure 3.4 - Q-factor in dB for transmission of 10.7 GBd DP-QPSK over 97 spans. Unfilled markers denote experimental data, while solid markers	

correspond to simulated results, and lines denote polynomial fits. CD only and 1 step per span are demonstrated.	69
Figure 3.5 - Variation of experimental Q-factor in dB with distribution of dispersion in Wiener-Hammerstein cascade compensation of 97 spans transmission DP-QPSK at 10.7 GBd. Shown for 0.5 dBm launch power and an 80.2 km step-size.	70
Figure 3.6 - Plot of improvement in inferred maximum Q-factor against mean dispersive block length for DP-QPSK at 10.7 GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.	72
Figure 3.7 - Plot of improvement in inferred optimum launch power against mean dispersive block length for DP-QPSK at 10.7 GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.	73
Figure 3.8 - Contour plot of experimentally determined Q-factor in dB against launch power and nonlinear step-size for Wiener cascade compensation of 20 spans transmission DP-QAM16 at 10.7 GBd. Nonlinear step-size of a single span lies at 80 km.	74
Figure 3.9 - Q-factor in dB for transmission of 10.7 GBd DP-QAM16 over 20 spans. Unfilled markers denote experimental data, while solid markers correspond to simulated results, and lines denote polynomial fits. CD only and 80 km nonlinear step-size are demonstrated.	75
Figure 3.10 - Plot of improvement in inferred maximum Q-factor against mean dispersive block length for DP-QAM16 at 10.7 GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.	76
Figure 3.11 - Plot of improvement in inferred optimum launch power against mean dispersive block length for DP-QAM16 at 10.7GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.	77
Figure 3.12 - Plot of improvement in inferred optimum launch power against number of nonlinear blocks for Wiener model nonlinearity compensation of DP-QPSK and DP-QAM16.	78
Figure 3.13 - Plot of improvement in inferred optimum launch power against nonlinear step size for Wiener model nonlinearity compensation of DP-QPSK and DP-QAM16.	78
Figure 3.14 - Relative increase in complexity for backpropagation of 10.7 GBd DP-QPSK and DP-QAM16	81

Figure 4.1 - Constellation diagrams for DP-QPSK and PS-QPSK in the phase space. Phase on the x polarization (ϕ_x) is plotted against phase on the y polarization (ϕ_y) (both in radians). Solid markers belong to both formats while hollow markers belong to PS-QPSK only.	86
Figure 4.2 - Ideal receiver sensitivity in symbol error rate against E_b/N_0 for DP-QPSK and PS-QPSK. For equivalent information carrying capacity, error rate of triplets of DP-QPSK is compared with error rate of quadruplets of PS-QPSK.	88
Figure 4.3 - 'Conventional' bit mapping for PS-QPSK after Karlsson & Agrell.	89
Figure 4.4 - 'Alternative' bit mapping scheme for PS-QPSK.	90
Figure 4.5 - Example of differential coding for PS-QPSK, based on 'alternate' bit mapping scheme.	91
Figure 4.6 - Ideal receiver sensitivity in bit error rate (BER) against E_b/N_0 for DP-QPSK and PS-QPSK with various coding schemes and bit mappings.	92
Figure 4.7 - Input polarization sensitivity of the DD-LMS with PS-QPSK modulation. Q-factor penalty in dB is plotted against the two angular parameters in the Jones matrix, with dark colour representing a low penalty and light colour representing a high penalty.	97
Figure 4.8 - Polarization sensitivity of the PS-CMA with and without initialisation algorithm. Mean Q-factor penalty is plotted against polarization angle. ...	98
Figure 4.9 - Constellation diagrams showing a single output polarization of PS-QPSK modulation when equalised with the PS-CMA equaliser (a); and PS-QPSK modulation when equalised with the DP-CMA equaliser (b). Noise loaded to 5.8 dB E_b/N_0 , for illustrative purposes.	99
Figure 4.10 - Performance of the PS-CMA with PS-QPSK modulation in the presence of PDL. Q-factor penalty in dB is plotted against the applied PDL in dB.	100
Figure 4.11 - Performance of the PS-CMA with conventionally coded PS-QPSK modulation in the presence of time varying polarization rotation.	101
Figure 4.12 - Performance of the PS-CMA with differentially coded PS-QPSK modulation in the presence of time varying polarization rotation.	102
Figure 4.13 - Q-factor penalty as a function of input polarization state for the equaliser algorithm as described in [14] (Johannisson et al.) compared with that described in section 4.4 (Millar et al.).	104

Figure 4.14 - Back-to-back comparison of required OSNR to achieve a BER of 10^{-3} for PS-QPSK and DOP-QPSK. Both systems are three channels at 112 Gb/s (7% FEC overhead) over 50 GHz WDM grid.	106
Figure 4.15 - Back-to-back comparison of required OSNR to achieve a BER of 2×10^{-2} for PS-QPSK and DP-QPSK. Both systems are three channels at 124.8 Gb/s (20% FEC overhead) over a 50 GHz WDM grid.....	107
Figure 5.1 - Constellation diagrams showing two orthogonal linear polarizations of an experimentally generated PS-QPSK signal.	110
Figure 5.2 - Experimental set-up to generate and transmit 42.9 Gb/s PS-QPSK (14.3 Gbaud) and DP-QPSK (10.725 Gbaud).	112
Figure 5.3 - Back-to-back measurements. Single-channel and WDM receiver OSNR sensitivity for (a) PS-QPSK and (b) DP-QPSK.....	115
Figure 5.4 - Transmission performance of 42.9 Gb/s PS-QPSK and DP-QPSK compared for a 7 channel WDM system on a 50 GHz frequency grid. Maximum reach is compared for a BER of 3.8×10^{-3}	116

1 Introduction

1.1 Introduction

The birth of optical fibre communications systems in the early 1970s produced an explosive growth in telecommunications with the introduction of low loss, high bandwidth silica fibres [1]. Increases in line rates in these early systems were achieved simply by increasing the modulation rate of on-off keyed (OOK) signals [2]. As symbol rates increased, systems were limited by the loss that could be tolerated before electrical regeneration was required.

At this stage, coherent detection was first investigated as a method of increasing the sensitivity of optical receivers [3]. While an improvement in sensitivity of up to 20 dB was achieved, coherent detection was superseded with the invention of the optical fibre amplifier [4].

Erbium doped fibre amplifiers (EDFAs) provide amplification in the optical domain, providing THz of gain bandwidth, enabling amplification of several wavelength channels. Further increases in reach became available with the invention of inverse dispersion fibres, allowing optical domain compensation of chromatic dispersion (CD).

While optical line rates continued to increase to tens of Gb/s, EDFAs enabled a large increase in total capacity per fibre. This increase was due to the use of several optical carriers of differing wavelengths on a single fibre, known as wavelength division multiplexing (WDM) [5]. The carriers are de-multiplexed at the receiver with a wavelength selective device such as an arrayed waveguide grating (AWG), and detected individually.

Although the increase in capacity enabled by EDFAs and WDM has scaled well in the past, a hard limit on capacity exists while OOK modulation is used, given that the maximum achievable information spectral density (ISD) is 1 b/s/Hz. During the development of 40 Gb/s transmission systems, the need to maintain compatibility with the ITU frequency grid of 50 GHz for WDM systems and reduce the need for

extremely high bandwidth electronic components became apparent. Due to these pressures, advanced modulation formats which are capable of transmitting more than one bit per symbol (and therefore an ISD of more than 1 b/s/Hz) became a highly desirable technical advancement [6].

Several transmission systems were implemented for 40 Gb/s transmission, two of which relied on quadrature phase shift keying (QPSK). Differential detection of differentially coded QPSK (DQPSK) provides a low optical complexity solution at 40 Gb/s, and relies on self-coherent detection using a delay line interferometer (DLI) and quadrature detection [7]. The delay line interferometer splits the incoming signal into two equal components, one of which is delayed by one symbol period. The two components are then recombined in a 2x2 coupler which has outputs in quadrature. These optical signals are then detected and sampled, giving the two bits of information which comprise the QPSK symbol.

Although DQPSK provided a relatively cost efficient method of achieving 40 Gb/s transmission over a 50 GHz WDM grid, self-coherent detection has attributes which limit the scaling of this technology to higher line rates. Due to the fact that the signal is mixed with itself, the improvement in sensitivity is less than that of full coherent detection. Signals are also limited by linear distortions such as polarization mode dispersion (PMD) and CD which scale linearly and with the square of baud rate respectively. Additional modulation density may be achieved only with the addition of further phase levels and additional DLI structures in the receiver, adding cost and complexity while significantly increasing complexity.

The second QPSK-based solution for 40 Gb/s transmission was fully coherent detection with digital post-processing of dual-polarization QPSK (DP-QPSK) [8]. Digital coherent detection involves splitting the incoming signal into two orthogonal linear polarizations, mixing the signal with a free-running local oscillator and performing detection of the in-phase and quadrature components for both polarizations. This scheme results in detection of all four dimensions of the optical field, thus preserving all information contained in the received signal [9]. After quantisation, distortions are removed from the signal with the use of DSP techniques, such that the channel is equalised, the polarization states are separated and the variation in phase and frequency between the local oscillator and source lasers is

removed. This scheme allows the use of polarization multiplexing without the need for adaptive optics, and also enables full compensation of arbitrary amounts of previously limiting effects such as PMD and CD [10]. The modulation format of choice for 40 Gb/s with coherent detection has become DP-QPSK, which carries 4 bits per symbol (resulting in an ultimate limit of 4 b/s/Hz ISD). The higher modulation density results in a lower bandwidth requirement for electrical components, with just 5 GHz required for most components along with 10 GSa/s for the analogue to digital convertors (ADCs).

While digital coherent detection requires additional optical components in comparison to other schemes (namely the optical hybrid, local oscillator and additional photo-detectors), and therefore incurs additional cost, both robustness to distortion and scaling to higher line-rates are greatly increased. This may be illustrated by the recent development in 100 Gb/s transmission systems. The commercial standard for 100 Gb/s transceivers [11] specifies the use of digital coherent detection in combination with DP-QPSK modulation. This increase in bit rate was enabled by simply increasing the bandwidth of electrical components and increasing the sampling rate of the ADCs to 56 GSa/s. The technical challenges involved in this scaling were almost entirely in the electrical domain. Previously limiting optical distortions such as PMD and CD may be fully compensated without penalty, leaving only nonlinearity as a limiting optical distortion.

The aim of this thesis is the development and analysis of digital signal processing techniques for enhancing performance in coherent optical communication systems. While much of the current digital techniques are identical to those found in the digital wireless literature, we will investigate improvements in performance to be obtained by tailoring modulation and DSP post-processing algorithms to the optical channel.

1.2 Chapter Overview

Chapter 1 gives a basic introduction to digital coherent detection and an overview of the aims, content and focus of this thesis.

Chapter 2 provides an overview of the literature and theoretical background regarding coherent detection, modulation formats, fibre transmission effects and impairments, digital post-processing algorithms for digital coherent detection and digital backpropagation.

Chapter 3 concerns an in-depth analysis of the performance of DP-QPSK and DP-QAM16 with and without nonlinearity compensation by digital backpropagation. The parameter space of the digital backpropagation algorithm is explored for a single channel transmission system such that an upper bound on the performance benefits due to backpropagation may be obtained.

Chapter 4 concerns the theoretical background and digital post-processing of polarization-switched QPSK modulation. This modulation format offers superior performance to conventional DP-QPSK modulation and may therefore be considered as an alternative to digital nonlinearity compensation in the low information spectral density regime.

Chapter 5 describes experimental generation and transmission of PS-QPSK signals. This experiment was the first reported long-haul WDM transmission of PS-QPSK, and was at the time the longest transmission distance reported for a 40G WDM system over a conventional non dispersion managed link.

Chapter 6 examines topics for future research which have been raised by the material in this thesis, and gives an overview of the research presented and conclusions drawn from the thesis.

1.3 List of Publications

The research presented in this thesis resulted in the following publications:

- (1) C. Behrens, D. Lavery, **D.S. Millar**, S. Makovejs, B.C. Thomsen, R.I. Killey, S.J. Savory and P. Bayvel, "Ultra-Long-Haul Transmission of 7x42.9 Gb/s PS-QPSK and PM-BPSK," *European Conference on Optical Communication (ECOC) 2011*, 2011, Mo.2.B.2. (Winner of the "ECOC Best Student Paper Presentation Award").
- (2) B.C. Thomsen, R. Maher, **D.S. Millar** and S.J. Savory, "Burst Mode Receiver for 112 Gb/s DP-QPSK," *European Conference on Optical Communication (ECOC) 2011*, 2011, Mo.2.A.5.
- (3) **D.S. Millar**, D. Lavery, S. Makovejs, C. Behrens, B.C. Thomsen, P. Bayvel, and S.J. Savory, "Generation and long-haul transmission of polarization-switched QPSK at 42.9 Gb/s," *Optics Express*, vol. 19, 2011, pp. 9296-9302.
- (4) **D.S. Millar** and S.J. Savory, "Blind adaptive equalization of polarization-switched QPSK modulation," *Optics Express*, vol. 19, 2011, pp. 8533-8538.
- (5) S. Makovejs, E. Torrenco, **D.S. Millar**, R.I. Killey, S.J. Savory, and P. Bayvel, "Comparison of Pulse Shapes in a 224Gbit/s (28Gbaud) PDM-QAM16 Long-Haul Transmission Experiment," *Optical Fiber Communication Conference (OFC/NFOEC) 2011*, 2011, OMR5.
- (6) E. Torrenco, S. Makovejs, **D.S. Millar**, I. Fatadin, R.I. Killey, S.J. Savory, and P. Bayvel, "Influence of Pulse Shape in 112-Gb/s WDM PDM-QPSK Transmission," *IEEE Photonics Technology Letters*, vol. 22, 2010, pp. 1714-1716.
- (7) **D.S. Millar**, S. Makovejs, I. Fatadin, R. Killey, P. Bayvel, and S. Savory, "Experimental characterisation of QAM16 at symbol rates up to 42 Gbaud," *European Conference on Optical Communication (ECOC) 2010*, 2010, P3.20.

- (8) **D.S. Millar**, S. Makovejs, C. Behrens, S. Hellerbrand, R.I. Killey, P. Bayvel, and S.J. Savory, "Mitigation of Fiber Nonlinearity Using a Digital Coherent Receiver," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, 2010, pp. 1217–1226.
- (9) **D.S. Millar** and S.J. Savory, "Nonlinear Digital Processing for Uncompensated Systems," *Signal Processing in Photonic Communications (SPPCom) 2010*, 2010, SPWB6.
- (10) S. Makovejs, **D.S. Millar**, D. Lavery, C. Behrens, R.I. Killey, S.J. Savory, and P. Bayvel, "Characterization of long-haul 112Gbit/s PDM-QAM-16 transmission with and without digital nonlinearity compensation," *Optics Express*, vol. 18, Jun. 2010, pp. 12939-47.
- (11) S. Makovejs, **D.S. Millar**, V. Mikhailov, G. Gavioli, R.I. Killey, S.J. Savory, and P. Bayvel, "Experimental investigation of PDM-QAM16 transmission at 112 Gbit/s over 2400 km," *Optical Fiber Communication Conference (OFC/NFOEC) 2010*, 2010, OMJ6 (shortlisted for the Corning Outstanding Student Paper Competition).
- (12) S. Makovejs, **D.S. Millar**, V. Mikhailov, G. Gavioli, R.I. Killey, S.J. Savory, and P. Bayvel, "Novel method of generating QAM-16 signals at 21.3 Gbaud and transmission over 480 km," *Photonics Technology Letters*, vol. 22, 2010, pp. 36–38.
- (13) P. Bayvel, C. Behrens, R. Killey, S. Makovejs, **D.S. Millar**, and S.J. Savory, "Coherent electronic compensation techniques for long-haul optical fibre transmission—opportunities and challenges," *European Conference on Optical Communication (ECOC) 2009*, 2009, 10.7.2.
- (14) **D.S. Millar**, S. Makovejs, V. Mikhailov, R. Killey, P. Bayvel, and S. Savory, "Experimental comparison of nonlinear compensation in long-haul PDM-

QPSK transmission at 42.7 and 85.4 Gb/s,” *European Conference on Optical Communication (ECOC) 2009*, 2009, 9.4.4.

1.4 Main Contributions

The main contributions to the literature of this thesis are as follows:

Chapter 3 describes the first experimental comparison of the efficacy of digital backpropagation for DP-QPSK and DP-QAM16, originally published in (6) [12]. This also provided the first thorough exploration of the parameter space for the backpropagation algorithm with experimental data. This work was performed in conjunction with S. Makovejs who assisted in performing measurements using the recirculating loop test-bed; C. Behrens and S. Hellerbrand who assisted with simulations of the transmission system used.

Chapter 4 describes a novel algorithm for the digital equalisation of PS-QPSK signals, first described in (2) [13]. This algorithm is used in conjunction with a filter initialization algorithm to achieve singularity-free convergence for up-to 5 dB of PDL. This was the second algorithm described for blind equalization of PS-QPSK (after [14], which was published while [13] was under review), and the first algorithm to achieve singularity-free blind equalization of this modulation format.

Chapter 5 describes the generation and long-haul transmission of PS-QPSK signals, and comparison to DP-QPSK systems, originally published in (1) [15]. This was the first demonstration of WDM transmission of PS-QPSK, the first long-haul transmission of PS-QPSK, and the highest reported reach for a 40G WDM transmission system using standard fibre and amplifiers. This work was performed with D. Lavery who assisted with performing the transmission measurements; S. Makovejs, C. Behrens and B.C. Thomsen who helped with the transmitter and WDM comb setup.

2 Literature Review and Theory

2.1 Abstract

In this chapter, the basic concepts of coherent detection are discussed. The details of polarization and phase diverse coherent detection are described for both single ended and balanced photo-detection. Local oscillator phase and frequency locking is studied, as are the constraints on analogue to digital conversion. Common modulation formats used for coherent detection are assessed and compared by noise sensitivity. Fibre transmission impairments are discussed, with different mathematical descriptions for fibre with and without chromatic dispersion, polarization mode dispersion, and Kerr nonlinearity. Noise resulting from amplified spontaneous emission in optical amplifiers is also described. Digital post-processing algorithms are then examined. These include: filters for the compensation of chromatic dispersion; adaptive equalization of PMD, polarization rotations and residual filtering; intradyne frequency offset estimation and compensation; and carrier phase estimation. The theoretical basis of nonlinearity compensation with digital backpropagation is then reviewed. Backpropagation is analysed using two and three block nonlinear models, and step sizes of more than and less than one span. Recent experimental results using digital backpropagation are noted, and alternatives to backpropagation discussed.

2.2 Introduction

From the inception of low-loss optical fibre in 1966 [1] until now, transmission systems have predominantly utilised intensity-modulation and direct-detection (IM-DD). These systems can be constructed with relatively simple optical components. For instance, the transmitter typically uses either a directly modulated laser [2], or a co-packaged laser and a single intensity modulator [16]. The receiver (neglecting any forward error correction (FEC) hardware) requires only a photodetector, a decision circuit and a clock recovery circuit. In these directly detected systems, only the intensity of the received field is recovered, and three of the four dimensions of the optical field are therefore discarded.

In the last few years, interest has been directed to coherent detection as an enabling technology for improving spectral efficiency [17] and mitigating fibre transmission impairments. Since all four dimensions of the optical field (amplitude and phase in two orthogonal polarizations) may be detected with a coherent receiver, digital signal processing may be utilised to mitigate optical impairments in the digital domain [18]. Coherent receivers may therefore leverage the rapid improvements in CMOS technology to provide highly robust and cost-effective transceivers.

Coherent optical receivers were initially investigated in the late 1980s as a serious commercial prospect [19]. Coherent detection offered a significant improvement in receiver sensitivity over direct detection systems of the time [20], in addition to frequency selectivity which enabled demultiplexing of WDM channels without the need for optical filters [21].

At approximately the same time, erbium doped fibre amplifiers (EDFAs) were developed [4]. EDFAs provide amplification in the optical domain, providing THz of gain bandwidth, therefore amplifying several wavelength channels [5]. These devices ameliorated the need for high sensitivity receivers and enabled long transmission distances without electronic regeneration. Further increases in reach became available with the invention of inverse dispersion fibres, allowing optical domain compensation of chromatic dispersion (CD).

While research into coherent detection continued, the scalability of IM-DD and WDM ensured that coherent systems were not commercially viable until recently.

The driving forces behind the resurgence of coherent detection may be considered as follows: system design constraints are such that significant amounts of existing fibre plant could not support 40 Gb/s direct-detection or DQPSK [22]; additionally, advances in CMOS technology have ensured that sampling rates in the region of 50 GSa/s and ASIC complexity in the region of 10^8 gates are commercially viable. Therefore we may consider that at 40 Gb/s coherent detection is feasible and desirable: at 100 Gb/s coherent detection is feasible and necessary.

2.3 Coherent Detection

Although differential detection enables greater receiver sensitivity than direct detection, it is still limited by linear distortions such as CD and PMD, and may not be used easily with polarization multiplexing. This form of coherent detection is also known as self-coherent or pseudo-coherent detection, and is generally used with simple receiver signal processing. A further problem with differential detection is that the signal is itself noisy, so the receiver sensitivity is reduced considerably.

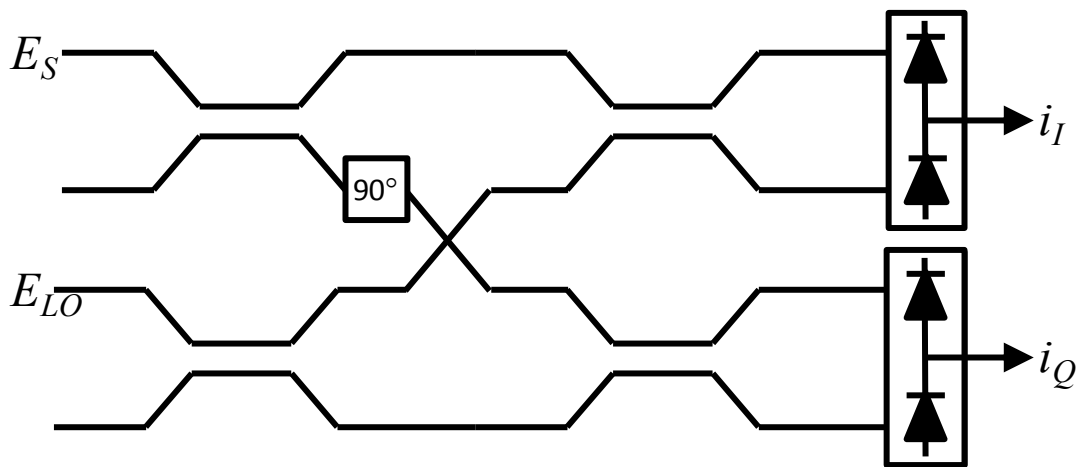


Figure 2.1 - Quadrature coherent detection

By employing full coherent detection (shown in Figure 2.1, with both quadratures detected), it is possible to gain 3 dB in receiver sensitivity [9], the additional spectral efficiency made available by polarization multiplexing and the ability to compensate for linear impairments to an arbitrarily high degree. As a result of these properties, for 100 Gb/s and above full coherent detection is the most attractive possibility.

2.3.1 Phase and Polarization Diverse Coherent Receiver

The phase and polarization diverse coherent receiver (shown in Figure 2.2) consists of three main stages: polarization splitting; phase diverse coupling and detection. In the first stage, the input signal and local oscillator are split into orthogonal polarizations by a pair of polarization beam splitters. A pair of polarization controllers are used to align the local oscillator polarizations with those from the input signal. This will ensure the maximum possible interference in the mixing stage. Signal – LO coupling is performed by a pair of 90° optical hybrids, which couple together local oscillator and signal for each polarization, and have a pair of outputs in quadrature. The four optical fields are then detected individually. Detection is normally performed with *P-I-N* photodiodes, either in single ended [23] or balanced configurations [24]. Photodiodes have an amplitude response of the form $I \propto |E|^2$, where I is the photocurrent and E is the incident electrical field [25]. It is the square-law response of the photodiodes which mixes the signal and LO fields together and enables coherent detection. When using single-ended detection, the local oscillator must be in the region of 20 dB higher than the signal input [26]. This effectively linearises the response of the photodiode by reducing the relative contribution of the direct-detection of the signal to the overall photocurrent. While this approach simplifies the receiver somewhat and reduces the component cost, the ratio of signal to LO power is a balance of penalties due to direct-detection of the signal and penalties due to the relative intensity noise (RIN) of the local oscillator.

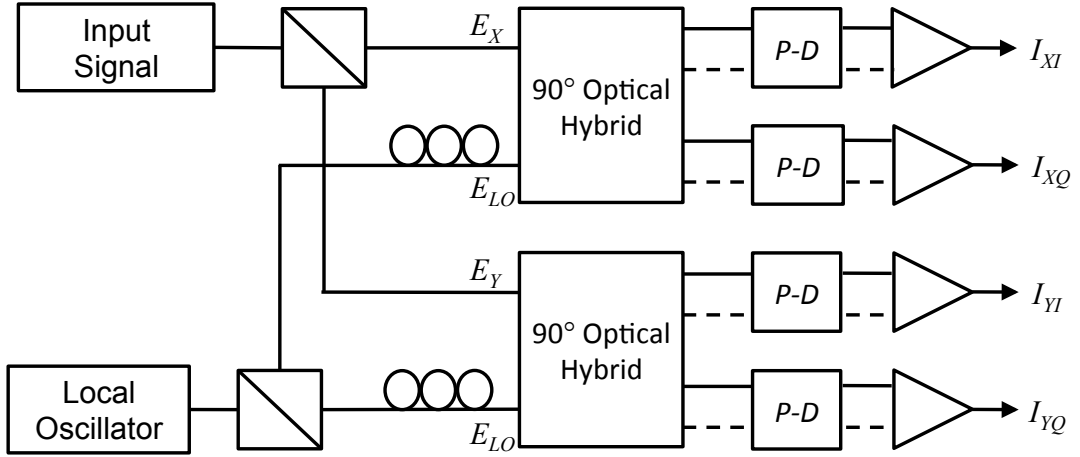


Figure 2.2 - Phase and polarization diverse coherent receiver. Dashed lines show signals required only when balanced detection is used.

The response of the phase and polarization diverse coherent receiver can be described as in (2.1) when used with single-ended detection [24] and symmetric 2x2 couplers for the optical hybrids.

$$\begin{pmatrix} I_{XI} \\ I_{XQ} \\ I_{YI} \\ I_{YQ} \end{pmatrix} \propto \begin{pmatrix} \text{Re}(E_X E_{LO}^*) \\ \text{Im}(E_X E_{LO}^*) \\ \text{Re}(E_Y E_{LO}^*) \\ \text{Im}(E_Y E_{LO}^*) \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 2|E_X|^2 + |E_{LO}|^2 \\ 2|E_X|^2 + |E_{LO}|^2 \\ 2|E_Y|^2 + |E_{LO}|^2 \\ 2|E_Y|^2 + |E_{LO}|^2 \end{pmatrix} \quad (2.1)$$

The first vector on the right hand side of the equation represents the part of the four photocurrents which are due to coherent detection. The second vector on the right hand side of the equation represents the current due to direct-detection. By making the local oscillator in the region of 20 dB more powerful than the signal, we can minimise the relative magnitude of the directly-detected signal terms (which are not constant power).

To overcome the constraints imposed on signal-LO power ratios, balanced photo-detection is often employed for coherent optical receivers. In this scenario, an 8 port optical hybrid is used, with a 180° phase shift between each quadrature pair. The pairs of outputs are then differentially amplified to eliminate the direct-detection components in the signal. The 8 output ports of the hybrid are given by (2.2):

$$\begin{pmatrix} I_{XI+} \\ I_{XI-} \\ I_{XQ+} \\ I_{XQ-} \\ I_{YI+} \\ I_{YI-} \\ I_{YQ+} \\ I_{YQ-} \end{pmatrix} \propto \begin{pmatrix} \frac{1}{2}\text{Re}(E_X E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ -\frac{1}{2}\text{Re}(E_X E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ \frac{1}{2}\text{Im}(E_X E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ -\frac{1}{2}\text{Im}(E_X E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ \frac{1}{2}\text{Re}(E_Y E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ -\frac{1}{2}\text{Re}(E_Y E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ \frac{1}{2}\text{Im}(E_Y E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \\ -\frac{1}{2}\text{Im}(E_Y E_{LO}^*) + \frac{1}{4}|E_X|^2 + \frac{1}{8}|E_{LO}|^2 \end{pmatrix} \quad (2.2)$$

After differential amplification, this becomes the 4-dimensional signal given by (2.3).

$$\begin{pmatrix} I_{XI} \\ I_{XQ} \\ I_{YI} \\ I_{YQ} \end{pmatrix} \propto \begin{pmatrix} \text{Re}(E_X E_{LO}^*) \\ \text{Im}(E_X E_{LO}^*) \\ \text{Re}(E_Y E_{LO}^*) \\ \text{Im}(E_Y E_{LO}^*) \end{pmatrix} \quad (2.3)$$

The received signal defined by (2.3) represents the ideal coherently received optical field, that is: no direct-detection terms, infinite common-mode rejection between differential pairs and perfectly matched optical path lengths in the hybrid resulting in an exact 90° difference between quadratures. Since this is the ideal case, in the simulations presented in this thesis, we will assume this model of the optical coherent receiver unless otherwise stated.

2.3.2 Local Oscillator Phase and Frequency Locking

During the mixing process, we are effectively downconverting the modulated carrier to a baseband equivalent, albeit one which is still modulated onto the intensity of an optical carrier (before photodetection). In a homodyne system, we would mix the signal with an LO with the same frequency and phase. This process requires locking the LO phase to the optical carrier, requiring some form of feedback loop and phase control. While some optical homodyne methods provide a relatively accurate and robust solution, all solutions of this type require a high degree of optical complexity, involving expensive and sensitive components. A more favourable method of

downconversion is to have a free-running local oscillator and recover the carrier frequency and phase digitally. When the intermediate frequency is less than the symbol rate, this is known as intradyning [27]. This allows lower optical complexity and the use of cheaper components, and can be combined with some form of coarse frequency control of the local oscillator to ensure that the system is able to lock continuously for many hours.

2.3.3 Analogue to Digital Conversion

After amplification, the four analogue electrical signals are digitised. For simplicity and performance in signal processing, it is convenient to have 2 samples per symbol. This rate of oversampling is not strictly necessary, but reduces the constraints on the required anti-aliasing filters [28] and enables compensation of a larger range of intradyne frequency offset. As these ADCs must operate in the region of 50 GSa/s, performance of these components is critical. Both the number of bits of resolution and timing jitter introduce uncertainty into the digitised signal. The jitter (or clock phase noise) reduces the accuracy of each sample, while the number of bits per sample determines the maximum accuracy of each sample. A metric of performance for ADCs is therefore the effective number of bits (ENOB), which accounts for both jitter and quantisation noise and models the effect of both as AWGN resulting in a noise floor of 6dB SNR per effective bit of resolution [29].

2.4 Modulation

2.4.1 Modulation Formats and Coding

A given modulation format proscribes an alphabet of possible ideal signal states to be transmitted through a communications channel. This signal alphabet can be considered as an N dimensional vector, where N is the order of the signal space. The N dimensional carrier is then modulated by the symbol vector. In the early days of optical communications, the only dimension modulated was intensity. While wireless digital communication has used multidimensional signalling for many years [30] the difficulties associated with detecting optical phase caused optical phase modulation to remain unexplored. Intensity modulation utilises a 1D symbol vector with M possible states representing $\log_2(M)$ bits of information per symbol. While early optical systems offered huge improvements in bandwidth and loss compared to

contemporaneous wireline and microwave systems, as only a single dimension of the signal was being modulated, the spectral efficiency of such systems was inherently low. With the invention of the erbium doped fibre amplifier (EDFA), it was no longer necessary to convert optical signals back into the electrical domain for amplification to achieve long distance transmission. This enabled a second dimension of the optical channel to be modulated: wavelength. While wavelength division multiplexing (WDM) enabled a vast increase in available transmission bandwidth, the manufacture of suitable optical components remained a constraint in using this dimension of the signal space freely. This large amount of bandwidth was revolutionary, but still had many constraints: linear distortions such as GVD and PMD were problematic and could be compensated only approximately in the optical domain, while envelope detection introduced nonlinearity while discarding most of the information about the incoming optical field. In this situation, designers began to explore ways of exploiting two further dimensions of the signal space: polarization and phase. Using polarization modulation with direct detection is somewhat difficult to implement, as stochastic polarization rotations within the channel must be tracked on a kHz time scale [31]. More success was had with self coherent detection using differential phase shift keying (DPSK) [32]. This uses a one symbol optical ‘delay and add’ line combined with mixing in the photodetectors to determine the difference in optical phase between successive symbols. While this proved a viable technology to implement 40 Gb/s links, a lack of flexibility and limitations by linear and nonlinear distortions makes this type of system undesirable for 100 Gb/s and beyond.

By moving to phase and polarization diverse coherent detection, the full optical field at the receiver can be detected. This enables compensation of linear effects such as chromatic dispersion (CD) and polarization mode dispersion (PMD) to be performed fully in the digital domain. While it is also possible to manipulate capacity in the frequency domain more fully using techniques such as orthogonal frequency division multiplexing (OFDM) [33], these systems introduce complexities and difficulties of their own. Here, as in the remainder of this thesis, we will concentrate on single carrier communication, where frequency domain multiplexing is performed on multiple separately modulated carriers, and we perform signal processing functions on each four dimensional carrier individually.

The question of how to best modulate in a four dimensional signal space remains an open one: although work has been done to determine the best formats in the noise and power limited case [34], this work is recent and the modulation formats proposed are largely untested. Despite this, the benefits of using four dimensionally optimised modulation formats may be significant, and this topic is discussed in detail in later chapters. Coherent optical systems generally use identical two dimensional modulation formats on each of the polarizations (a technique known as polarization multiplexing). The more common of these two dimensional formats are described below.

2.4.2 Phase Shift Keying

Phase shift keying (PSK) encodes information onto the phase of the carrier. This can provide high spectral density for low symbol rates. The disadvantage of modulation formats of this kind is that the tolerance to both phase noise and AWGN is greatly reduced for high orders. Due to these limitations, M-PSK is widely used for only M of 2 and 4 (Figure 2.3).

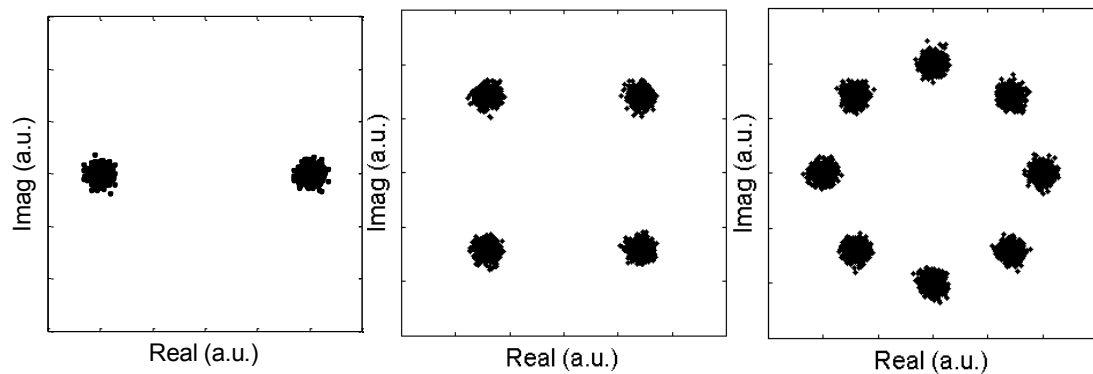


Figure 2.3 - Constellation diagrams for BPSK (left), QPSK (centre) and 8PSK (right). Noise loaded to E_s/N_0 of 24.5 dB.

2.4.3 Quadrature Amplitude Modulation

Quadrature amplitude modulation (QAM) refers to simultaneous amplitude modulation of two carriers of the same frequency which are in quadrature (often denoted in-phase (I) and quadrature (Q)). These modulation formats are often represented as phasors, with the real and imaginary parts of the phasor representing I and Q respectively. The most common of these types of formats results when the I and Q components are both modulated with several equally spaced amplitudes with zero mean (Figure 2.4). The resulting format is often called square QAM for obvious

reasons. For equal spectral density to PSK, QAM has better tolerance to AWGN [35]. Although phase recovery becomes more complex, phase margin for QAM is also better than PSK of equivalent order. Another form of QAM is formed by multiple rings of PSK with different amplitudes. This can be implemented with aligned rings (star QAM, see Figure 2.5, left), which has somewhat reduced SNR tolerance, or with adjacent rings offset (star QAM, see Figure 2.5, right), which provides better SNR tolerance but has no Gray code [36]. This modulation format has become relatively popular in coherent optical communication and is often referred to simply as 8-QAM [37].

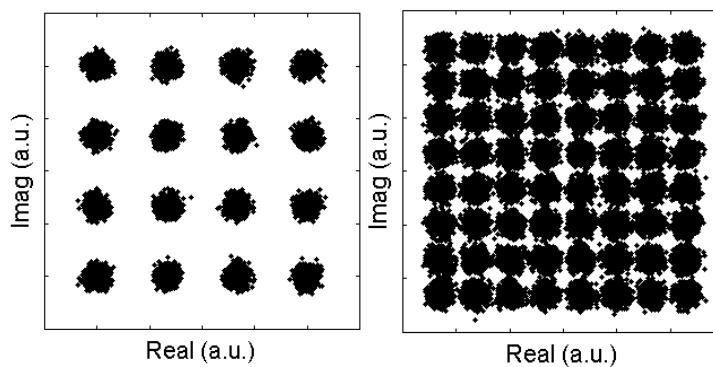


Figure 2.4 - Square QAM16 (left), and QAM64 (right). Noise loaded to E_s/N_0 of 24.5 dB.

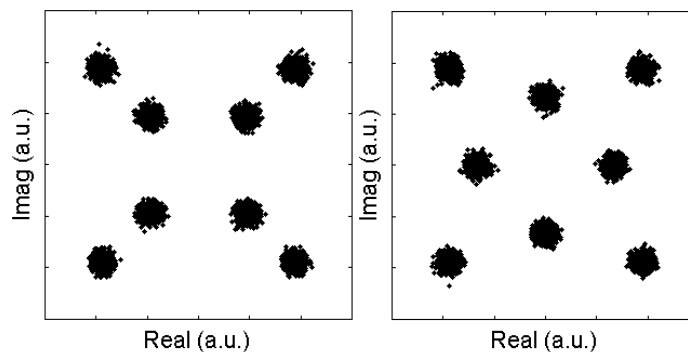


Figure 2.5 - Star QAM8 (left), and offset star QAM8 (right). Noise loaded to E_s/N_0 of 24.5 dB.

2.4.4 Bit-to-Symbol Mapping

Once a modulation format has been selected, it is necessary to decide how to map bits of information onto the signal space. By ensuring that adjacent constellation points in the signal space also have minimum binary Hamming distance, it can be shown that BER is minimised for a given SNR. The most well known code for this

application is binary Gray code [35] (see Table 1). This is a cyclic code with minimum Hamming distance and can be used for any square QAM and PSK.

Decimal	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
Hexadecimal	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Binary	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
Gray Coded Binary	0000	0001	0011	0010	0110	0111	0101	0100	1100	1101	1111	1110	1010	1011	1001	1000

Table 1 - Gray coding and its relation to other numbering systems

The advantage of Gray coding is illustrated below in Figure 2.6. While the Gray coded constellation always has a single bit difference between adjacent points, the standard binary code has some adjacent points which have 2 bits different. This will lead to a penalty in error rate, compared with Gray coding.

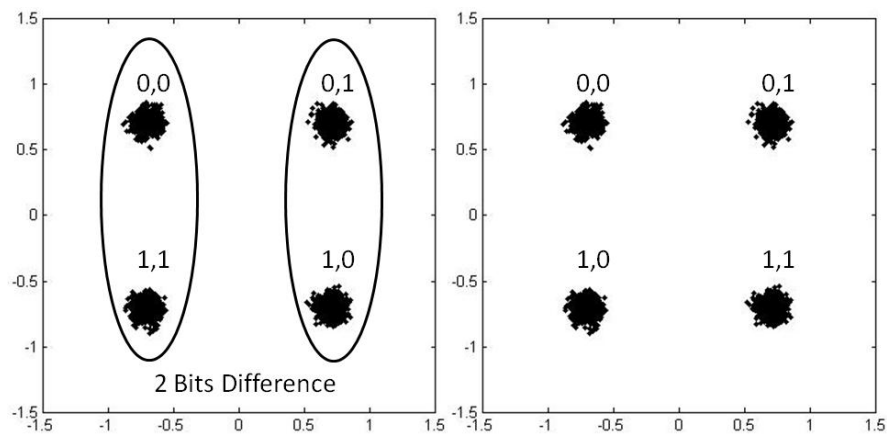


Figure 2.6 - QPSK constellations with standard binary coding (left), and Gray coding (right).

While this coding scheme is known to be optimal when it exists, there are some constellations for which it is impossible to construct, such as offset star QAM.

In addition to Gray coding, for applications where phase noise is a significant impairment, it may be desirable to sacrifice some of the error rate performance relating to noise, by differentially encoding the signal with respect to phase. In the noise limited case, differential coding will approximately double the bit error rate, but will also mitigate the effects of errors in phase recovery known as cycle slips. As

the phase estimation algorithms used in digital communications systems are typically bounded on some phase interval (for instance, the interval $\{-\pi, \pi\}$), this bounded phase estimate must be unwrapped such that the unbounded phase estimate may be applied to the recovered signal. Errors in this unwrapping process may result in some phase ambiguity, as the constellation becomes rotated relative to its position before the cycle slip occurred. These errors may result in long bursts of errors, which could be a source of system outage. Since for deployed systems reliability is of paramount importance, this method of coding is of great interest.

Differential coding may be described with extremely simple mathematics, described below in (2.4). These equations describe the relation between coded bits y and uncoded bits x at time index i , for each bit p in the symbol [38].

$$\begin{aligned} y_{i,p} &= y_{i-1,p} \oplus x_{i,p} \text{ for encoding} \\ x_{i,p} &= y_{i-1,p} \oplus y_{i,p} \text{ for decoding} \end{aligned} \tag{2.4}$$

However, differential phase coding is not fully compatible with Gray coding for all modulation formats. For instance, since square QAM is rotationally invariant for phase rotations which are multiples of $\pi/2$, we desire differential encoding for the two bits which represent the quadrant. This differential encoding will remove the possibility of burst errors due to cycle slips. The bits which encode the position within the quadrant must be rotationally invariant however, and this makes perfect Gray coding impossible.

2.4.5 Comparison of the Performance of Modulation Formats

It is possible to derive theoretical limits on the performance of different modulation formats based upon the use of optimal transmitters and receivers, and performance being purely impaired by additive white Gaussian noise [35]. Modulation formats which are currently considered feasible are shown below in Figure 2.7 and Table 2 whether considered desirable or not.

These asymptotes for optimal receiver performance assume a two dimensional channel and additive White Gaussian noise as the only impairment. The transmitter and receiver are assumed to have ideal matched filters, providing a signal that is band-limited at the Nyquist frequency and is ISI free.

The noise level is described here using the signal processing convention of E_b/N_0 , where E_b is the mean energy per transmitted bit and N_0 is the mean noise energy per symbol. This metric enables comparison between modulation formats with differing cardinality at identical bit rates, as the SNR is normalised to the number of bits per symbol of the modulation format (unlike, for example, E_s/N_0). If we convert this to the optical convention of optical signal to noise ratio over a bandwidth of 12.5 GHz, we may get an indication of how they might perform in, for example, a 112 Gb/s system of the kind required for the 100 GbE standard.

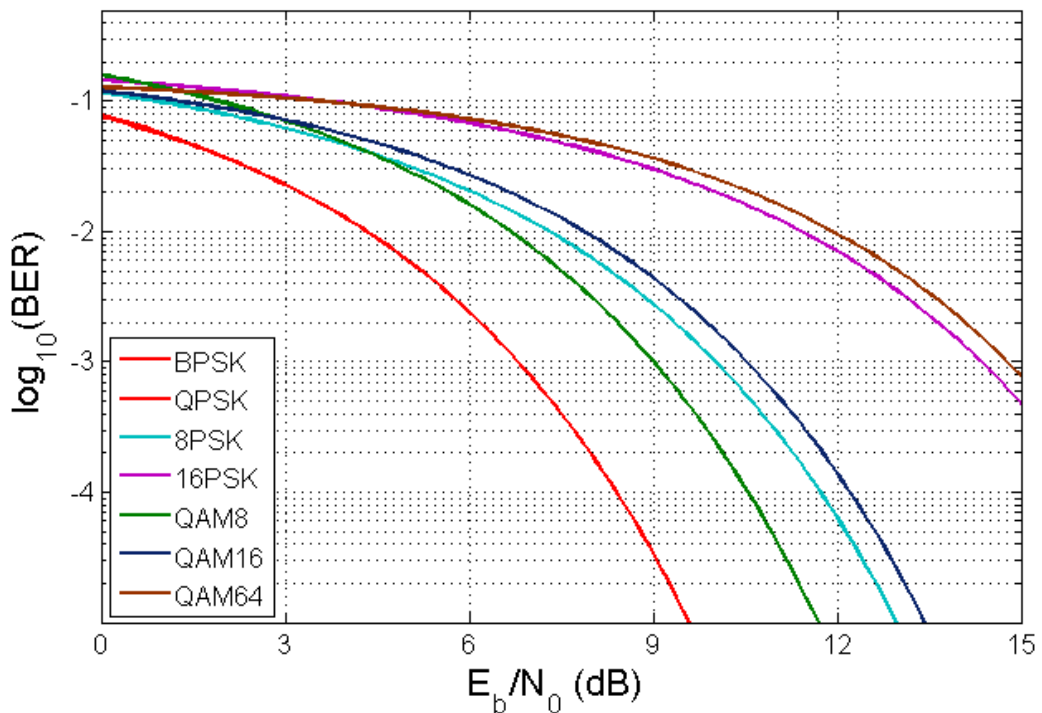


Figure 2.7 - Noise limited receiver performance for various modulation formats.

While the optimal performance asymptotes are derived for a two-dimensional channel, generalisation to the case of dual-polarization modulation formats is quite simple.

Modulation Format	Symbol Rate (GBd)	Possible Spectral Efficiency (b/s/Hz)	Required OSNR @BER=10 ⁻³ (dB/0.1nm)
DP-BPSK	56	2	13.3
DP-QPSK	28	4	13.3
DP-8PSK	18.7	6	16.5
DP-16PSK	14	8	20.8
DP-QAM8	18.7	6	15.5
DP-QAM16	14	8	17.0
DP-QAM64	7	10	21.2

Table 2 - Required OSNR for different modulation formats at a bit rate of 112 Gb/s and BER of 10⁻³

We can clearly see from Table 2 that the required OSNR becomes larger with increased spectral density. It is also clear from these figures that the preferred formats for 6 and 8 bits per symbol are respectively DP-QAM8 and DP-QAM16.

2.4.6 Optical Modulators

An important component in any optical transmitter is the optical modulator which converts the analogue electrical baseband signal into the carrier band of hundreds of THz. While there are many ways to modulate an optical carrier, many devices are unsuitable for advanced modulation formats. The most common and important modulator for coherent systems is the triple Mach-Zehnder modulator.

2.4.6.1 Mach-Zehnder Modulators

The Mach-Zehnder interferometer is a structure which simply splits a beam into two parts, shifts the phase of one with respect to the other and then recombines them [39].

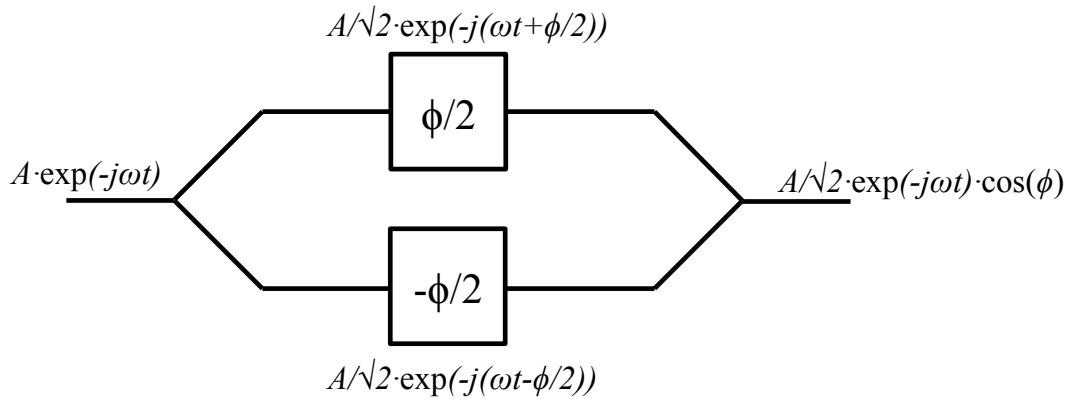


Figure 2.8 - Push-pull MZM operation.

By inserting an electro-optic modulator (EOM) into each of the arms of an MZI, it is possible to control the relative phase of the two arms, and therefore the intensity of the interference products at the output. This device is known as a Mach-Zehnder modulator (MZM) (Figure 2.8).

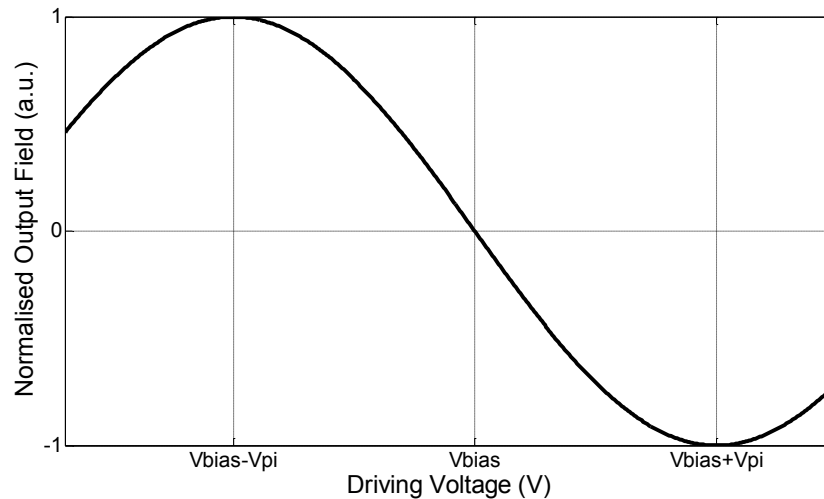


Figure 2.9 - MZM transfer function.

The transfer function of the MZM therefore has the form of a sinusoid (Figure 2.9). As the transfer characteristic of the modulator is inherently nonlinear, the output optical signal to noise ratio (OSNR) is a function of both the input electrical SNR and the range over which the modulator is driven. For modulation formats such as BPSK and QPSK, only the most extreme swing of the modulator is needed. By biasing the modulator to its null point and driving it over $2V_{pi}$, it is possible to suppress the electrical noise on the driving signal. This is due to the fact that the modulator is at its most nonlinear around V_{pi} , small changes in the driving voltage

have relatively little influence on the output field. Conversely, around the null point the transfer function of the MZM is approximately linear. This results in the noise on the electrical driving signal being transferred into the optical domain. While this effect may be suppressed when using a format such as BPSK or QPSK, when using higher order QAM or PSK it is desirable to use only the approximately linear part of the transfer function. While this is a method makes the implementation of highly complex modulation formats a good deal simpler, there is inevitably a reduction in transmitted SNR.

While this device gives good control of the output field, it will only modulate a single axis of the complex plane. To gain full control of the output field, a second MZM in quadrature is used, and combined with the first by way of a third MZI. This device will modulate the carrier over the full complex plane, with the two nested MZMs modulating I and Q separately. This modulator is known by many names, including: triple Mach-Zehnder modulator, I-Q modulator, Cartesian modulator or nested Mach-Zehnder modulator.

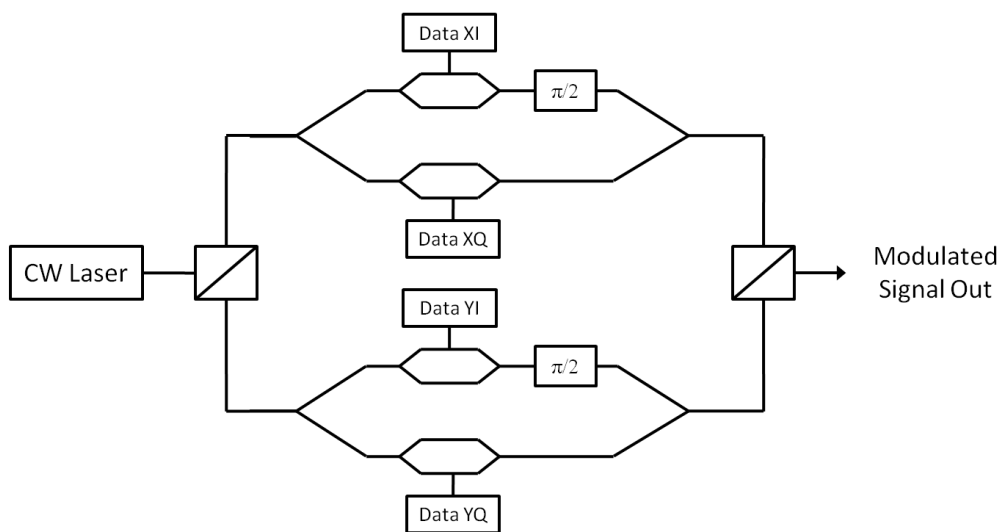


Figure 2.10 - Dual-polarization triple Mach-Zehnder modulator.

For generating dual-polarization modulation formats, typically two triple Mach-Zehnder modulators are used in parallel, each modulating an orthogonal polarization (Figure 2.10). The two unmodulated carriers come from the same laser and are split into orthogonal linear polarizations with a polarization beam splitter (PBS), before the two single polarization modulated signals are multiplexed together with a second PBS.

2.5 Fibre Transmission Impairments

The three most important fibre transmission impairments in communications systems are chromatic dispersion (CD), polarization mode dispersion (PMD) and the Kerr nonlinearity. The first two of these are linear, and may be compensated to arbitrary levels with the use of linear DSP.

2.5.1 Chromatic Dispersion

Chromatic dispersion occurs due to the differing group velocities of wavelengths in optical fibre. There are two sources of this dispersion: waveguide geometry (resulting in waveguide dispersion) and the properties of Silica (resulting in material dispersion). Here, we will consider the net dispersion in the fibre (the combination of waveguide and material dispersion).

The response of a purely dispersive optical fibre in the retarded time frame is described by (2.5) [40]:

$$\frac{\partial \mathbf{E}}{\partial z} = \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} \mathbf{E}; \quad (2.5)$$
$$\beta_2 = \frac{\partial^2 \beta}{\partial \omega^2}$$

where \mathbf{E} is the electrical field which varies in time t along a spatial dimension z . The fibre propagation constant β determines the group velocity and therefore dispersion of the fibre. Chromatic dispersion is normally defined in terms of the dispersion parameter D , which is given by (2.6):

$$D = -\frac{2\pi c\beta_2}{\lambda} \quad (2.6)$$

The dispersion parameter D has units of ps/nm/km, and is approximately 17 ps/nm/km for standard single mode fibre at 1550 nm.

2.5.2 Polarization Mode Dispersion

Single mode fibre has a single propagational mode, which consists of two spatial modes, which may be considered to represent two orthogonal polarization states. In a perfectly symmetric fibre, these modes are identical and have indistinguishable

propagation characteristics. However, imperfections in the waveguide geometry combined with mechanical stress introduced by vibrations and variations in temperature along the fibre result in birefringence. In a birefringent material, refractive index (and therefore group velocity) is determined by the axis along which the light is polarized. The difference between these fast and slow axes will determine the polarization mode dispersion over a short length of fibre. This model is complicated somewhat over a significant distance of fibre, since the fibre may be considered to be composed of many birefringent sections (commonly ~ 0.1 km for SSMF), which all have arbitrary alignments of their fast and slow axes. Since these axes have arbitrary alignment, PMD scales with the square root of distance [41] and has a Maxwellian distribution, leading to units of ps/ $\sqrt{\text{km}}$.

2.5.3 The Kerr Nonlinearity

The Kerr nonlinearity is the dominant nonlinear effect for telecommunication systems over silica fibre. The effect results in the refractive index of the fibre varying with the square of the instantaneous electrical field intensity [40]. This effect results in nonlinear phase modulation, which is converted by chromatic dispersion to amplitude modulation resulting in distortion in both phase and amplitude. The exact and numerical solutions of the nonlinear Schrödinger equation are discussed in detail in section 2.7. Other nonlinear effects occur in optical fibre, such as Raman and Brillouin scattering. These effects are due to the nonlinear interaction of photons and vibrational quasi-particles known as phonons. While Raman and Brillouin scattering are important in some optical fibre applications (such as Raman amplification), they are generally insignificant for conventional coherent systems in the $1.5 \mu\text{m}$ band, and are normally neglected in nonlinear fibre simulations.

2.5.4 Additive Noise

The dominant source of additive noise in coherent transmission systems is amplified spontaneous emission (ASE). This noise results from spontaneous emission within EDFAs, which is then subsequently amplified. As the noise is approximately white and Gaussian over the signal bandwidth, it is often described as additive White Gaussian noise (AWGN).

2.5.5 Laser Phase Noise

Laser phase varies largely according to the fundamental mode of the laser cavity. While lasers are amongst the highest quality resonators which can be made, imperfections in the evolution of phase are important when the phase is used for carrying information. The phase of a laser may be described as a ‘random walk’, where phase change between two points in time has a Gaussian distribution where the variance is proportional to the time between observations. Laser phase noise is therefore modelled as a Wiener process with a variance given by (2.7) [9]:

$$\sigma^2 = 2\pi\Delta\nu T \quad (2.7)$$

where T is the time between observations and $\Delta\nu$ is the 3 dB optical linewidth. As high-level modulation formats with little phase margin have become more common, phase noise has become a critical parameter with typical values of laser linewidth being between 100 kHz and 1 MHz.

2.6 Digital Signal Processing Algorithms for Coherent Detection

After detection, the received signals are digitised, and then processed to compensate for distortion and impairments. While none of the functionalities provided by DSP are strictly necessary for some kind of transmission to be achieved (although not used when coherent detection was proposed in the 1980s [3], [42], sophisticated DSP was considered an important advantage in the resurgence of coherent optical communications in the 2000s [43], [44]), the increase in performance and the reduction in constraints for the design of systems provided by DSP ensure that they are an extremely attractive proposition. The signal flow model of the DSP used in the coherent communication systems in this thesis is provided in Figure 2.11. It should be noted however, that as all of these functional blocks other than the decision circuit are linear, the order of the blocks may be interchanged, and this ordering is simply one possible implementation.

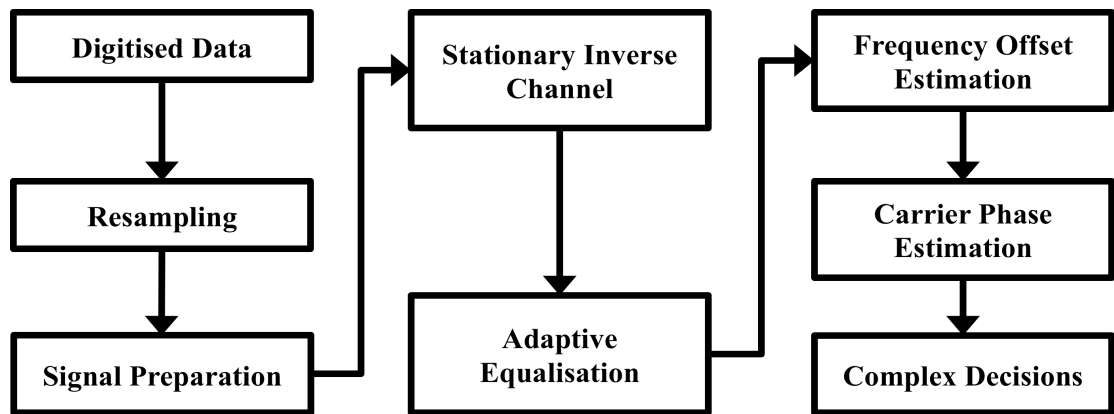


Figure 2.11 - DSP Signal Flow Model

2.6.1 Resampling

When asynchronous sampling is used, the digitised data is first resampled to a rate of two samples per symbol. When offline processing is used, this may be performed with a relatively simple upsample-filter-decimate algorithm. As this requires *a-priori* knowledge of the symbol rate, this is not an approach that is possible to implement in hardware. The most commonly used method of digital clock recovery is the use of an interpolating filter which is sampled at a rate determined by a nonlinear timing error detector such as the Gardner loop [45].

2.6.2 Signal Preparation

The signal is then prepared for processing in a number of ways. The four signals must be de-skewed to compensate for the relative differences in optical path lengths inside the receiver. Any DC component of the signals must be removed to compensate for signal components which are artefacts of the receiver structure and not present in the optical spectrum. The signals are then normalised individually, such that each polarization has unit mean power. This simplifies much of the signal processing performed later and ensures that different electrical powers (due to differences in the gain of the photodetectors and amplifiers used for each of the four signals) are compensated for.

2.6.3 Stationary Inverse Channel

Due to the ability of coherent receivers to capture the full optical field, impairments due to the nature of the channel may be compensated. A major linear impairment due to the nature of optical fibre is chromatic dispersion (CD), whereby different

wavelengths of light have different group velocities in the fibre. This effect may be considered as the combination of bulk material dispersion which is due to the properties of the silica which is used to make fibres, and waveguide dispersion which is due to the geometry of the fibre waveguide. As this effect is stationary, we can simply perform linear filtering with the inverse transfer function of that of the fibre. This filtering may be performed either in the frequency domain, or the time domain.

2.6.3.1 Frequency Domain Dispersion Compensation

In order to compensate for chromatic dispersion (CD), the linear part of the nonlinear Schrödinger equation may be solved, yielding a frequency domain response given by (2.8) [40]:

$$H(z, \omega) = \exp\left(-j \frac{D\lambda_0^2 z}{4\pi c} \Delta\omega^2\right) \quad (2.8)$$

where λ_0 is the reference wavelength where D is known, and $\Delta\omega$ is the difference in natural frequency between the wavelength of interest and λ_0 . By calculating the FFT of the signal and multiplying the signal spectrum (per polarization) with the conjugate of this response, CD may be fully compensated.

2.6.3.2 Time Domain Dispersion Compensation

While the method described above provides an accurate method for the compensation of CD, it may not be desirable to have to calculate the FFT of the signal (which is computationally intensive) in order to compensate for the effects of CD. A filter design which may be implemented in the time domain may therefore be useful, particularly as an FIR filter which can be made inherently stable. The inverse Fourier transform of (2.8) gives the impulse response of the linearised filter, which is a rotating vector with linearly increasing frequency. By truncating this response at the Nyquist frequency (half of the sampling frequency), both aliasing and non-causality may be avoided [10]. The design of the FIR filter is then as follows in (2.9):

$$\begin{aligned}
N &= 2 \cdot \left\lceil \frac{|D|\lambda_0^2 z}{2cT^2} \right\rceil + 1 \\
-\left\lfloor \frac{N}{2} \right\rfloor &\leq k \leq \left\lfloor \frac{N}{2} \right\rfloor \\
a_k &= \sqrt{\frac{jcT^2}{D\lambda_0^2 z}} \exp\left(-j \frac{\pi cT^2}{D\lambda_0^2 z} k^2\right)
\end{aligned} \tag{2.9}$$

here, N is the number of taps in the filter; k is the tap index; a_k the tap coefficients; D is the dispersion coefficient of the fibre; λ_0 is the reference wavelength; z is the distance of transmission; T is the sampling period; c is the speed of light. This filter design enables the compensation of arbitrarily large amounts of dispersion, although with small amounts of dispersion, pass-band ripple due to the small number of taps may be a significant impairment.

2.6.4 Adaptive Equalization Algorithms

Adaptive equalisation in coherent communications is most often performed with adaptive FIR filtering. These filters have been developed from the pioneering work of Bussgang [46] and Sato [47]. These equalisers are desirable as they are relatively simple to design, robust, and may be used with complex baseband signals by making the filter coefficients and signals complex. The most commonly used class of equaliser – the Bussgang equaliser - has an essential structure as described in Figure 2.12. It consists of three major parts: an FIR filter to implement the inverse channel, a memoryless non-linearity which estimates how far the filter is from the desired response, and an update algorithm which determines new filter coefficients based upon the old coefficients, the estimated position on the error surface, and the input signal to the filter [46].

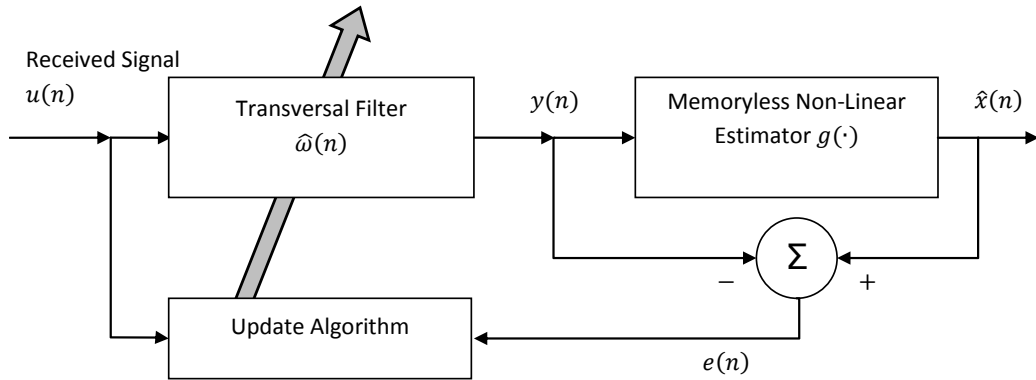


Figure 2.12 - Block diagram of a Bussgang equaliser.

2.6.4.1 MIMO Processing with Bussgang Equalisers

With coherent polarization multiplexed communication, the equaliser structure used to separate the two incoming polarizations (which are constantly rotating stochastically on the Poincaré sphere) is very similar to that used in MIMO wireless systems. The filter structure used is known as a four-filter butterfly structure (Figure 2.13). Each output is an arbitrary combination of the input signals, thus enabling both deconvolution of the signal from the channel, and separation of the two source signals.

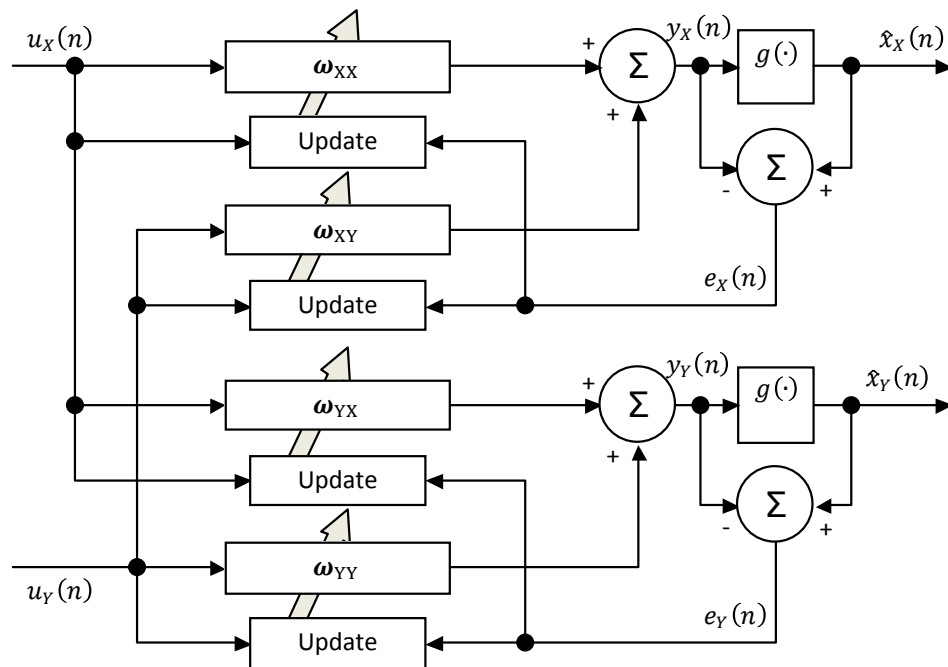


Figure 2.13 - The 2x2 MIMO Bussgang Equaliser, figure taken from [46].

2.6.5 Least Mean Squares Algorithm

The least mean squares (LMS) algorithm is an algorithm which trains the filter coefficients. There are many of these adaptation algorithms, such as: least squares, least mean squares, recursive least squares and the Kalman filter [46]. While all of these learning algorithms could be used with coherent optical systems, the LMS algorithm [48] has been adopted near universally. The LMS algorithm has attractive properties for high speed communications, it is highly stable, may converge rapidly, and most importantly requires relatively little computational effort when compared to algorithms such as the RLS algorithm or the Kalman filter.

The algorithm adapts the filter coefficients based upon the derivative of the cost function with respect to the filter coefficients. The cost function describes the deviation of the filter output from the desired response.

For a single-input, single-output (SISO) complex baseband channel, the filter is adapted according to (2.10):

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n - \mu \hat{\nabla} J_n \quad (2.10)$$

where $\hat{\mathbf{w}}_n$ is the tap vector at instant n , and $\hat{\nabla} J_n$ is the estimated gradient of the cost surface with respect to $\hat{\mathbf{w}}_n$.

The adaptation of the filter is often described by an error term e_n given by (2.11):

$$e_n = -\frac{\mathbf{u}_n v_n^*}{\hat{\nabla} J_n};$$

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mu e_n \mathbf{u}_n v_n^*; \quad (2.11)$$

where: $v_n = \hat{\mathbf{w}}_n^H \mathbf{u}_n$

Here, \mathbf{u}_n is the input vector at instant n ; v_n is the instantaneous output of the equaliser; $*$ is the complex conjugate; and H represents the Hermitian conjugate.

The adaptation of a 2x2 MIMO filter using the least mean squares algorithm given by (2.12):

$$\begin{aligned} \mathbf{h}_{xx} &= \mathbf{h}_{xx} + \mu e_x \mathbf{x}_{in} x_{out}^*; & \mathbf{h}_{xy} &= \mathbf{h}_{xy} + \mu e_x \mathbf{y}_{in} x_{out}^*; \\ \mathbf{h}_{yx} &= \mathbf{h}_{yx} + \mu e_y \mathbf{x}_{in} y_{out}^*; & \mathbf{h}_{yy} &= \mathbf{h}_{yy} + \mu e_y \mathbf{y}_{in} y_{out}^* \end{aligned} \quad (2.12)$$

where \mathbf{x}_{in} and \mathbf{y}_{in} are the input vectors to the equaliser on the x and y polarizations respectively; and the output of the equaliser is given by (2.13):

$$x_{out} = \mathbf{h}_{xx}^H \mathbf{x}_{in} + \mathbf{h}_{xy}^H \mathbf{y}_{in}; \quad y_{out} = \mathbf{h}_{yx}^H \mathbf{x}_{in} + \mathbf{h}_{yy}^H \mathbf{y}_{in} \quad (2.13)$$

2.6.5.1 Decision-Directed Equalization

The decision-directed error function makes a symbol estimation decision on the received signal (which in the case of QPSK is the complex signum function), and subtracts the received symbol from the decision. The cost surface is given by (2.14):

$$J = \langle (D(x_{out}) - x_{out})^2 \rangle; \quad (2.14)$$

This may be differentiated with respect to the tap vector, leading to the decision-directed error-function, which is given by (2.15):

$$e_x = D(x_{out}) - x_{out}; \quad (2.15)$$

where e_x is the error term; x_{out} is the filter output; $\langle \cdot \rangle$ represents the expectation operator; and D is the hard decision function. This decision function for DP-QPSK modulation is given by (2.16):

$$D(x_{out}) = 1/\sqrt{2} \left(\text{sgn}(\text{Re}(x_{out})) + j \text{sgn}(\text{Im}(x_{out})) \right); \quad (2.16)$$

where $\text{sgn}(\cdot)$ is the signum function. While this is a simple error function to implement, its efficiency depends on making a series of successive correct decisions. For this reason decision-directed equalisers are often started with another algorithm to converge to the vicinity of a minimum before being switched to decision-directed mode [49].

2.6.5.2 The Godard Algorithm

The Godard algorithm attempts to filter the incoming signal to one of constant power, other than the influence of noise [50]. This makes the algorithm ideal for modulation formats such as PSK, which does have constant power. It can also be

used for formats such as QAM, although for higher orders the variation in power between different symbols can be so great that the high steady state value of the error function makes accurate convergence impossible. The Godard algorithm cost function is given by (2.17):

$$J = \langle (R_p - |x_{out}|^p)^2 \rangle; \quad (2.17)$$

This may be differentiated, leading to the error function given by (2.18):

$$e_x = x_{out} |x_{out}|^{p-2} (R_p - |x_{out}|^p); \quad (2.18)$$

where p is the order of the algorithm and R_p is the modulus which the equaliser attempts to approach. The most popular implementation of the Godard algorithm is known as the constant modulus algorithm (CMA). This is a Godard equaliser with $p=2$, and is normally used with a radius of 1 for simplicity of calculation (2.19):

$$e_x = x_{out} (1 - |x_{out}|^2); \quad (2.19)$$

While the Godard algorithm is somewhat limited by the modulation formats for which it has good performance, it is extraordinarily robust. This robustness means that it can be used for pre-convergence of an equaliser before switching to decision-directed mode, which displays better SNR tolerance. The ability to separate the functions of equalisation and phase and frequency recovery in this family of equalisers has proven very attractive to those working in coherent optical communication, and adaptations of this algorithm have been made to provide this functionality for modulation formats such as QAM where the ‘multi-modulus’ nature of the modulation format impacts upon the efficacy of the equaliser [37], [51], [52]. The most popular adaptation is the radially directed equaliser (RDE), which performs an initial decision as to which is the modulus of the incoming signal, and then uses the CMA error function as described in (2.19) with the appropriate value of R to determine the updated filter coefficients.

2.6.6 Frequency Offset Compensation Algorithms

When using an intradyne coherent receiver with a free running local oscillator, the transmitter and receiver lasers are not frequency locked. This results in some residual frequency offset in the received signal. For this offset to be compensated, the

intradynic frequency must be estimated and compensated to translate the received signal into the baseband. There are two main classes of algorithms suitable for this application: stochastic accumulators which perform some manner of time domain averaging of the differential carrier phase per symbol, and spectral methods which calculate the frequency offset using a Fourier transform of the signal over some time period. In order to estimate the frequency offset, it is necessary to remove variations in phase due to modulation. A simple method for removing the effects of modulation is to look at the difference between the complex decision of a received symbol and the symbol itself. A more elegant method for removing the effects of modulation is to use an M^{th} power nonlinearity [53] for a modulation format with M degrees of rotational symmetry.

Two methods for generating an estimate of frequency offset in the time domain are described here (for the sake of simplicity using QPSK). They are the block window accumulator (BWA), and the gliding window accumulator (GWA).

2.6.6.1 **Block and Gliding Window Accumulators**

These algorithms calculate an estimate of phase per symbol and average over many symbols to remove the influence of noise [54]. For the example of QPSK with a 4th power nonlinearity to remove the phase modulation, the block window estimation algorithm is described by (2.20):

$$\Delta\phi_k = \arg\left(\sum_{i=2}^k (x_i x_{i-1}^*)^4\right)/4; \quad (2.20)$$

$$y_k = x_k \exp(-jk\Delta\phi_k)$$

where * represents the complex conjugate; x is the sequence of input symbols; y is the sequence of output symbols; k is the time index; and $\Delta\phi_k$ is the phase offset per symbol estimated at time k . This algorithm provides a sufficient quality of estimate while the frequency offset remains constant. Unfortunately, this algorithm does not track variations in frequency offset well, as it estimates the IF over all previous received symbols. In order that frequency offset may be tracked, a gliding window is often used. This is described by (2.21)

$$\Delta\phi_k = \arg\left(\sum_{i=k-L}^{k+L} (x_i x_{i-1}^*)^4\right)/4;$$

$$y_k = x_k \exp(-jk\Delta\phi_k)$$
(2.21)

where L is the feedforward delay of the estimator.

2.6.6.2 Frequency Domain Estimation

Frequency offset may be estimated in the frequency domain by using an M^{th} power nonlinearity to remove modulation, and then finding the maximum power in the Fourier transform of the resultant signal. This peak will be located at the offset frequency, as shown in (2.22) where T_s is the symbol period and FFT denotes the discrete Fourier transform.

$$C(f) = \text{FFT}(x^4);$$

$$C(f_0) = \max(C(f));$$

$$\Delta\phi = \frac{2\pi f_0}{T_s};$$

$$y_k = x_k \exp(-jk\Delta\phi)$$
(2.22)

While computationally intensive, this method is accurate without requiring training or convergence, and may be used for short signals such as those captured in experiments where the sequence length may be on the order of 10^5 symbols [53].

2.6.7 Carrier Phase Estimation Algorithms

Carrier phase estimation is necessary due to the varying relative phase of the Tx and LO lasers, due to their non-zero laser linewidth. As with frequency estimation, modulation must be removed from the signal, and noise must be averaged out. A somewhat unique aspect of coherent optical systems is that the relative linewidth is much higher than wireless systems. This means that phase estimation is much more of a limiting problem for optical systems, and these algorithms are consequently much more important to system performance than in wireless. We will discuss the Viterbi and Viterbi algorithm and digital phase-locked loop which is often used for higher order QAM modulation.

2.6.7.1 Viterbi and Viterbi algorithm

By far the most commonly used algorithm for PSK based systems is the Viterbi and Viterbi algorithm [55]. The algorithm uses an M^h power nonlinearity to remove the modulation from the signal. The instantaneous phase estimate is given by (2.23):

$$\hat{\phi}_k = \arg \left(\sum_{i=k-L}^{k+L} F(\rho_i) x_i^M \right) / M; \quad (2.23)$$

where F is an arbitrary nonlinear function, M is the degree of rotational symmetry in the constellation, and ρ_i is the magnitude of x_i . Most often, F is simply unity, and the only weighting of symbols by magnitude is done by the M^h power. This phase estimate is bounded between $-\pi/M$ and π/M , while the phase difference between the two lasers is bounded on $-\pi, \pi$. This initial estimate of carrier phase must therefore be unwrapped in order that ‘cycle-slips’ may be avoided. The unwrapping and phase correction algorithm is given by (2.24).

$$\begin{aligned} a_k &= a_{k-1} + \left[1/2 - \frac{M}{2\pi} (\hat{\phi}_k - \hat{\phi}_{k-1}) \right]; \\ \phi_k &= \hat{\phi}_k + a \frac{2\pi}{M}; \\ y_k &= x_k e^{-j\phi_k} \end{aligned} \quad (2.24)$$

‘Cycle-slipping’ occurs when the true phase crosses a boundary while the phase estimate is not unwrapped, or conversely when the true phase does not cross a boundary while the phase estimate is unwrapped. This results in a rotation of the signal constellation and subsequent mis-detection of the signal phase, resulting in burst errors. Differential coding can mitigate the worst burst errors which could cause signal failure, however, unwrapping is still highly desirable to improve performance.

2.6.7.2 Decision-Directed Phase-Locked Loop

While the Viterbi and Viterbi algorithm provides good performance for MPSK and DP-MPSK modulation, it is challenging to adapt for other modulation formats. One such modulation format which has garnered significant interest recently is DP-QAM16. Although the Viterbi and Viterbi algorithm may be used [56], [57] some constellation points must be discarded so that all points have one of four phases. To counteract this, a generic algorithm which uses all points in the signal constellation is

desirable. An algorithm of this description which has been often employed in the literature [12], [58-60] is the decision-directed digital phase-locked loop (DD-PLL).

The DD-PLL uses a per-symbol decision-directed differential phase estimate and an exponentially weighted integration to provide an unwrapped phase estimate [49] (2.25):

$$\begin{aligned} y_k &= x_k e^{-j\phi_k}; \\ e_k &= y_k - D(y_k); \\ \phi_{k+1} &= \phi_k - \mu \text{Im}(y_k - e_k^*) \end{aligned} \tag{2.25}$$

where μ is the constant of integration. This algorithm is very useful in offline experiments, as it may be used for any modulation format and is well described in the literature. However, the algorithm is biased toward previous estimates of phase (and is therefore not an ML estimator) and also requires feedback at the symbol rate which is not practical in a highly parallelised DSP ASIC. To solve the problem of the estimator bias, an alternative algorithm was developed in [61]. This algorithm was not used for the results discussed in this thesis, and will therefore not be discussed further.

2.7 Digital Backpropagation

2.7.1 Nonlinear Channel Models

Digital backpropagation is a method of nonlinearity compensation which has generated much interest recently [62-65]. It exploits the knowledge of the physical behaviour of the optical fibre as a nonlinear channel, by approximating the inverse nonlinear channel, most commonly described by the nonlinear Schrödinger equation (NLSE). The solution of the NLSE is approximated by the split-step Fourier method (SSFM), commonly used to simulate nonlinear transmission in optical fibres. While this algorithm may be used to approximate an inverse channel, the channel is inherently limited by additive White Gaussian noise (AWGN) which stems from amplified spontaneous emission (ASE) due to inline optical amplification. This noise cannot be removed by any filter due to its random nature, and will additionally result in nonlinear phase noise (due to the Gordon-Mollenauer effect) which occurs due to the nonlinear interaction of the signal and noise.

When performing simulations of fibre transmission systems, it is conventional to use the coupled polarization nonlinear Schrödinger equation (CP-NLSE) [40] (2.26).

$$\begin{aligned}\frac{\partial}{\partial z} E_X &= -\frac{\alpha}{2} E_X + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} E_X - j\gamma \left(|E_X|^2 + \frac{2}{3} |E_Y|^2 \right) E_X - \frac{j\gamma}{3} E_X^* E_Y^2 \\ \frac{\partial}{\partial z} E_Y &= -\frac{\alpha}{2} E_Y + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} E_Y - j\gamma \left(|E_Y|^2 + \frac{2}{3} |E_X|^2 \right) E_Y - \frac{j\gamma}{3} E_Y^* E_X^2\end{aligned}\quad (2.26)$$

For digital backpropagation, it is common to use the Manakov equation (2.27) [66] without PMD for the inverse channel model.

$$\begin{aligned}\frac{\partial}{\partial z} E_X &= -\frac{\alpha}{2} E_X + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} E_X - j\gamma \frac{8}{9} (|E_X|^2 + |E_Y|^2) E_X \\ \frac{\partial}{\partial z} E_Y &= -\frac{\alpha}{2} E_Y + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} E_Y - j\gamma \frac{8}{9} (|E_Y|^2 + |E_X|^2) E_Y\end{aligned}\quad (2.27)$$

PMD is neglected for the inverse channel model due to its random, stochastic nature. The Manakov equation is more computationally efficient than the CP-NLSE for long distance simulations, where total transmission distance is greater than 1000 km [67]. This model accounts for the residual birefringence of the fibre, and the effect that this has on the state of polarization and nonlinearity within the fibre. Since the residual birefringence scatters the signal state of polarization on a significantly smaller scale than the fibre nonlinear length, fibre nonlinearity acts on both polarizations equally as described by the right-most term in (2.27). Here, $\mathbf{E}=[E_X, E_Y]$ is the optical field in delayed time, subscripts x and y denote orthogonal linear polarization states, α is the fibre loss parameter, β_2 the group velocity dispersion parameter and γ the nonlinearity parameter. Though the Manakov equation is widely used for simulation of fibres with residual birefringence, application of this equation for digital backpropagation algorithms has only recently been made in [68] and formalised in [69].

Equation (2.27) may be split into its constituent linear and nonlinear parts, resulting in:

$$\begin{aligned}\frac{\partial \mathbf{E}}{\partial z} &= (\widehat{D} + \widehat{N})\mathbf{E}, \quad \text{where} \\ \widehat{D} &= \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2}, \\ \widehat{N} &= -j\gamma \frac{8}{9} \mathbf{E}^H \mathbf{E} - \frac{\alpha}{2}, \text{ and} \\ \mathbf{E} &= [E_X \quad E_Y]^T\end{aligned}\tag{2.28}$$

From (2.28), we may derive an exact solution of the Manakov equation, which forms the basis of the split-step type solutions (2.29):

$$\mathbf{E}(z + h, T) = \exp\left(h(\widehat{D} + \widehat{N})\right) \mathbf{E}(z, T)\tag{2.29}$$

where h is the nonlinear step-size.

2.7.2 Split-Step Methods

The following approximation is central to all split-step Fourier method numerical solutions to both the CP-NLSE (2.26), and the Manakov equation (2.27) which we will investigate here. We may say that for sufficiently small step h :

$$\exp\left(h(\widehat{D} + \widehat{N})\right) \mathbf{E}(z, T) \approx \exp(h\widehat{D})\exp(h\widehat{N})\mathbf{E}(z, T)\tag{2.30}$$

By approximating the non-commutable operators \widehat{D} and \widehat{N} as commutable [70], we form an approximate solution to the NLSE which may be evaluated analytically. While the approximation in (2.30) is very basic, this approximation of the exact solution in (2.29) forms the basis of all split-step type solutions, and is a good approximation over a sufficiently small step-size. A common refinement to (2.30) is to evaluate the nonlinear part of the solution with a constant envelope profile and varying intensity. This modification allows larger step sizes to be used as the solution does not imply constant power throughout the step as (2.30) does.

By assuming that the only change in the electric field over the nonlinear step is loss (and consequently that dispersion may be considered to act separately), we may normalise the solution of the nonlinear part of the Manakov equation to the varying power profile within the step and remove the loss term.

$$\begin{aligned}
\widehat{N}'(\mathbf{E}', z') &= \int_z^{z+h} \widehat{N}(\mathbf{E}', z') dz' \\
&= L_{Eff} \widehat{N}(\mathbf{E}, z'), \text{ where} \\
\mathbf{E}'(z+h, T) &= \exp(-\alpha h/2) \mathbf{E}(z, T)
\end{aligned} \tag{2.31}$$

The approximation in (2.31) gives us a nonlinear step which includes loss and the total nonlinear phase shift over the spatial step. This is effectively a multiplication by the effective nonlinear length L_{Eff} , given by (2.32) [40]:

$$L_{Eff} = \frac{1 - \exp(-\alpha h)}{\alpha} \tag{2.32}$$

The accuracy of this solution to the Manakov equation can be improved by applying the dispersion operator in two equal parts, before and after the nonlinear operator. This leads to the symmetric split-step Fourier method [40] defined by (2.33).

$$\begin{aligned}
&\exp\left(h(\widehat{D} + \widehat{N})\right) \mathbf{E}(z, T) \approx \\
&\exp\left(h/2 \widehat{D}\right) \exp(L_{Eff} \widehat{N}) \exp\left(h/2 \widehat{D}\right) \mathbf{E}(z, T)
\end{aligned} \tag{2.33}$$

2.7.3 Two- and Three- Block Nonlinear Models

The two variations of the split-step method may be described in terms of two and three block nonlinear models from nonlinear systems theory [71], for each short length of fibre over which a split-step is taken. Nonlinear models of particular interest are: the Wiener model which consists of a linear block followed by a memoryless nonlinear block (Figure 2.14); the Hammerstein model, consisting of a memoryless nonlinear block followed by a linear block (Figure 2.15); and the Wiener-Hammerstein model, which represents the concatenation of the Wiener and Hammerstein models, that is, a linear block followed by a memoryless nonlinear block, followed by a second linear block (Figure 2.16).

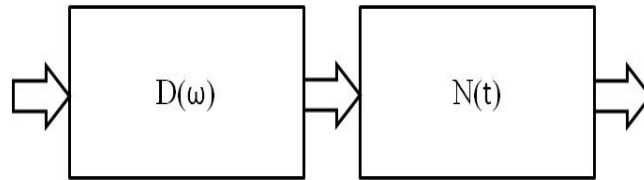


Figure 2.14 - The Wiener model

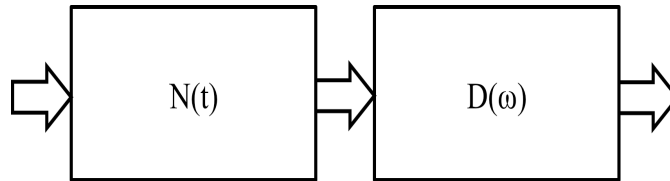


Figure 2.15 - The Hammerstein model

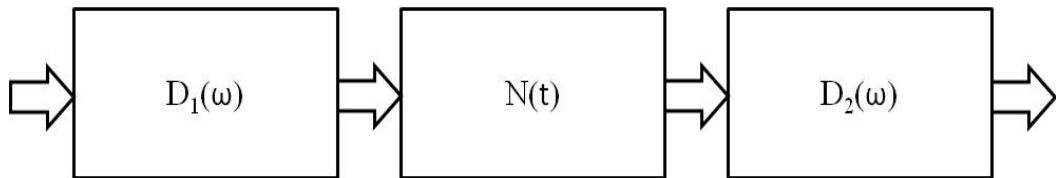


Figure 2.16 - The Wiener-Hammerstein model

It is immediately apparent that the Wiener-Hammerstein model (Figure 2.16) is used to represent the behaviour of an incremental section of fibre in the symmetric split-step method. These three block systems are then cascaded to form an approximation of the entire channel. Similarly, the bulk-step approach refers to a cascaded two block (Hammerstein model) approximation of the channel. The difference between the two and three block nonlinear models stems largely from their relative accuracy, which is analysed in the next subsection. It should be noted that while the bulk-step approximation of the forward channel corresponds to a cascade of Hammerstein systems, when this model is used for compensation (that is, approximation of the inverse channel) the order of the blocks is reversed, and the bulk-step model therefore refers to a cascade of Wiener systems.

2.7.4 Channel Models and Accuracy Considerations

The improvement in accuracy that the symmetric split-step method offers over the ‘bulk’ step method given in (2.30) may be quantified by examining the error term generated by the approximations in (2.30) and (2.33), by use of the Baker-Campbell-Hausdorff formula [70] for the commutability of operators (2.34). This formula gives us an analytical insight into both the relative accuracy of the symmetric-step and

bulk-step variants of the nonlinear channel model, and the effects of step size and channel model on the accuracy of the digital backpropagation algorithm.

$$\exp(h\widehat{D})\exp(h\widehat{N}) = \exp\left(h\widehat{D} + h\widehat{N} + h^2/2(\widehat{D}\widehat{N} - \widehat{N}\widehat{D}) + \dots\right) \quad (2.34)$$

In the Baker-Campbell-Hausdorff formula (2.34), for convenience we have included the dominant (first) term only of the error series. This is equivalent to the error in the bulk-step approximation given in (2.30). The error term resulting from the symmetric split-step method may be described as follows (2.35):

$$\begin{aligned} \exp(h/2\widehat{D})\exp(h\widehat{N})\exp(h/2\widehat{D}) &= \exp(h/2\widehat{D})\exp(\Theta) \\ &\text{where:} \\ \Theta &= h\widehat{N} + h/2\widehat{D} + h^2/2\left[\widehat{N}, \frac{\widehat{D}}{2}\right] + \exp(h/2\widehat{D})\exp(\Theta) \\ &= \exp\left(h\widehat{D} + h\widehat{N} + h^3/6\left[\widehat{N} + \frac{\widehat{D}}{2}, \left[\widehat{N}, \frac{\widehat{D}}{2}\right]\right] + \dots\right) \end{aligned} \quad (2.35)$$

We note from (2.34) that the dominant error term is proportional to h^2 , while in (2.35) the dominant term is proportional to h^3 . This indicates (a fact well known by those familiar with optical fibre simulations using the split-step Fourier method) that the symmetric split step method will give greater accuracy with an identical spatial step size.

An essential factor in the application of this solution, however, is that the solution of the Manakov equation is to be performed at the receiver with a noisy signal and DSP. It is therefore in our interest to perform not the most accurate reverse propagation, but the least complex with sufficient accuracy. Much recent research [26], [68], [72] has demonstrated that for receiver based digital backpropagation, a single nonlinear step per span with the bulk-step method may be considered a sufficiently accurate solution of the Manakov equation.

2.7.5 Digital Backpropagation Algorithms

The coherent receiver-based nonlinearity-compensation algorithms used in this investigation may be categorised as two variations of digital backpropagation. These

are the Wiener and Wiener-Hammerstein inverse-channel approximations (corresponding to the bulk-step and symmetric-step methods); each of which may be applied to step sizes of less than one span, or one span or more.

Dispersive steps may be performed in the frequency domain, using the dispersion operator defined in (2.28). This method describes a circular convolution of the dispersion operator (2.28) and the time domain signal, and is given in (2.36), where \mathcal{F} represents the Fourier transform.

$$\exp(h\widehat{D}) \mathbf{E} = \mathcal{F}^{-1} \left\{ \exp(h\mathcal{F}\{\widehat{D}\}) \mathcal{F}\{\mathbf{E}\} \right\} \quad (2.36)$$

Nonlinear operations are performed in the time domain with the nonlinear operator normalised to both nonlinear step-size and launch power. For nonlinear step sizes of a single span or greater, the nonlinear operator is that defined in (2.37).

$$\widehat{N}(t, z_{NL}) = j\varphi z_{NL} (|E_X(t)|^2 + |E_Y(t)|^2) P_L \quad (2.37)$$

Here P_L represents launch power in mW, z_{NL} is the nonlinear step size, t represents the retarded time frame and φ is the nonlinear phase shifting coefficient, which is a constant to be optimised. All other symbols retain their conventional meaning or those defined previously.

In the case of more than one nonlinear step per span, the nonlinear operator is modified to account for the exponentially varying power profile within the span. This leads to a nonlinear operator as defined in (2.38):

$$\widehat{N}(t, z_{NL}) = j\varphi 10^{\left(\frac{sL}{10n}\right)} z_{NL} (|E_X(t)|^2 + |E_Y(t)|^2) P_L \quad (2.38)$$

where n is the number of steps per span, s is the index of the step within the span and L is the fibre loss per span in dB.

2.7.6 Recent Experimental Results in Nonlinear Backpropagation

Since coherent detection has re-emerged a critical field of research, a significant amount of research effort has been directed at experimental demonstration of digital nonlinearity compensation. The most important of these results are presented in Table 3:

Year	Modulation Format	Bit Rate per Channel	Number of Channels	Channel Spacing	Performance Improvement
2009 [68]	DP-QPSK	112Gb/s	20	100GHz	2 dBQ
2009 [64]	DP-QPSK	112Gb/s	72	50GHz	0.25 dBQ
2009 [69]*	DP-BPSK	12Gb/s	3	7GHz	16 dBQ
2009 [26]	DP-QPSK	42.7Gb/s, 85.4Gb/s	1	-	33-50% reach
2009 [71]	DP-QPSK	112Gb/s	40	100GHz	1.7 dBQ
2009 [73]	DP-Co-OFDM	111Gb/s	1	-	13% reach
2010 [74]	DP-Co-OFDM	61.7Gb/s	1	-	2.2 dBQ
2010 [12]	DP-QAM16	112Gb/s	1	-	67% reach
2010 [75]	DP-QPSK	112Gb/s	1,10	50GHz, 100GHz	3.2-46% reach
2010 [12]	DP-QPSK, DP-QAM16	42.7Gb/s, 85.4Gb/s	1	-	1 dBQ
2010 [76]	DP-Co-OFDM	224Gb/s	7	50GHz	0.5 dBQ
2011 [77]	DP-Co-OFDM	448Gb/s	3	80GHz	0.5 dBQ
2011 [78]	DP-QAM16	224Gb/s	1,3	50GHz	12-19% reach
2011 [79]	DP-QPSK	112Gb/s	1,40	100GHz	1.5-2.5 dBQ
2011 [80]**	DP-QPSK	112Gb/s	80	50GHz	1 dBQ
2011 [81]	DP-Co-OFDM	112Gb/s	1,10	50GHz	0.5-1.4 dBQ

Table 3 - Recent experimental results from nonlinear digital backpropagation (* all three channels backpropagated; ** hybrid system with 79 10.7 Gb/s NRZ neighbours). Publications resulting from work in this thesis are in bold.

As can be seen from Table 3, when a single channel nonlinear backpropagation technique is employed (i.e. excluding the low baud rate WDM experiment presented in [69]), the increase in performance is limited to approximately 2.5 dBQ (where dBQ indicates decibels of Q-factor), or a 67% increase in reach. In a comparative

study between DP-QPSK and DP-QAM16 operating over the same system [12] it was found that when systems are operated near their linear limit, the potential benefit from mitigating intra-channel nonlinearities is marginally greater for DP-QAM16 than DP-QPSK. This study also revealed, while the maximum benefit in Q-factor was similar, the increase in optimum launch power was significantly higher (2.5 dB versus 1.8 dB) reflecting the more significant increase in launch power dynamic range. To date, much of the work in nonlinear digital backpropagation has focused on using a single nonlinear step per span. This approach has been demonstrated to achieve significant improvements in performance for single channel [26], [61], [72], [74] and widely spaced WDM systems [68], [73]. However, improvement in performance for WDM systems operating on a dense grid with high spectral efficiency (2 b/s/Hz or higher) has so far been small [64], [75-78], [80], [81]. In order to recover some of this benefit, one approach is to reduce the step size in order to determine if this affords any improvement. While some work has shown that this approach is unnecessary for reverse propagation of a single channel [12], significant benefits have been demonstrated when performing reverse propagation of several received WDM channels as a single signal [69], [82].

2.7.7 Alternative Approaches to Nonlinearity Compensation

Although digital backpropagation has become the dominant method used in research into fibre nonlinearity compensation, many other methods have been discussed in the literature. While the optimal nonlinear decoder (known as the Viterbi decoder or maximum likelihood sequence estimator (MLSE)) is well known and widely used in wireless communications [83], it results in unacceptable levels of computational complexity for coherent optical systems. Complexity is proportional to M^m , where M is the symbol alphabet and m is the channel memory [84]. While the Viterbi algorithm was utilised for IM-DD systems where modulation was binary and channel memory was low [85], it is prohibitively complex for coherent systems with no optical dispersion management and high-level modulation, and has so far only been used in a sparing manner [84].

Other statistically based nonlinear algorithms such as *maximum-a-posteriori* (MAP) detection have been used successfully in simulation and lab-based transmission experiments [86], [87]. While performance improvements obtained are notable, these

algorithms also have the problem of extremely high complexity which scales exponentially with channel memory and signal alphabet.

Volterra models of fibre nonlinearity have been utilised for some time [88], [89], and have recently been utilised in the design of nonlinear equalisers [90-92]. The popularity of these equalisers has again been limited by their computational complexity.

A further widely investigated technique is digital pre-compensation of nonlinearity using transmitter based DSP and a digital-to-optical transmitter [93-95]. Compensating nonlinearity at the transmitter is attractive due to the potential to simplify processing by the use of nonlinear look-up tables to implement processing of the undistorted data [96]. Despite this, the requirements of *a-priori* knowledge of the optical channel and a high-speed DAC at the transmitter have restricted interest in this technology for line rates beyond 10G.

2.8 Summary

In this chapter, the concepts of coherent detection were discussed. Polarization and phase diverse coherent detection was described with both single ended and balanced photo-detection. Local oscillator phase and frequency locking was examined, along with constraints on analogue to digital conversion. Modulation formats commonly used for coherent detection were introduced and compared by noise sensitivity. Fibre transmission impairments were then discussed, with different mathematical descriptions for fibre with and without chromatic dispersion, polarization mode dispersion, and Kerr nonlinearity. Noise resulting from amplified spontaneous emission in optical amplifiers was also detailed. Digital post-processing algorithms discussed include: filters for the compensation of chromatic dispersion; adaptive equalization of PMD, polarization rotations and residual filtering; intradyne frequency offset estimation and compensation; and carrier phase estimation. The theoretical basis of nonlinearity compensation with digital backpropagation was then elaborated, with models for backpropagation using two or three block nonlinear models, for step sizes of more than and less than one span. Previous progress in digital backpropagation was noted, and other techniques for digital nonlinearity compensation were described.

In the next chapter, we examine the benefits in performance available with nonlinearity compensation when the digital backpropagation algorithm is used. When digital backpropagation is applied to a single channel transmission system we are able to improve performance in the high-power nonlinear regime. The impact of step-size in the backpropagation algorithm is investigated in detail and the relationship between performance improvement and computational complexity is investigated.

3 Experimental Analysis of Digital Backpropagation

3.1 Abstract

Coherent detection with receiver based digital signal processing (DSP) has recently enabled the mitigation of fibre nonlinear effects. We investigate the performance benefits available from the backpropagation algorithm for dual-polarization quadrature phase shift keying (DP-QPSK) and 16-state quadrature amplitude modulation (DP-QAM16). The performance of the receiver using a digital backpropagation algorithm with varying nonlinear step size is characterized to determine an upper bound on the suppression of intra-channel nonlinearities in a single-channel system. The results show that for the system under investigation DP-QPSK and DP-QAM16 have maximum step sizes for optimal performance of 160 km and 80 km respectively. Whilst the optimal launch power is increased by 2 dB and 2.5 dB for DP-QPSK and DP-QAM16 respectively, the Q-factor is correspondingly increased by 1.6 dB and 1 dB, highlighting the importance of studying nonlinear compensation for higher level modulation formats.

3.2 Introduction

Exponential growth in capacity requirements in recent years has led to rapid improvements in the spectral efficiency of optical communications systems [97]. While this growth was previously sustained by the introduction of wavelength division multiplexing (WDM), with on-off keying (OOK) systems, this approach yields a theoretical maximum of 1bit/s/Hz over the bandwidth of the optical channel [98]. By utilizing coherent detection with phase and polarization diversity, it becomes possible to detect the full four dimensional signal space of amplitude and phase on two orthogonal polarizations rather than the single dimension of total power used with direct-detection (DD). As all four dimensions of the optical field are detected by a phase and polarization diverse coherent receiver, they may all be used for modulation, improving the achievable spectral efficiency. The detection of all four dimensions of the optical field also enables the equalization of previously limiting linear transmission impairments such as chromatic dispersion (CD) and polarization mode dispersion (PMD). With the elimination of linear transmission impairments, attention has turned to the mitigation of nonlinear impairments, which digital coherent receivers cannot completely compensate, stimulating research into nonlinearity mitigating receiver subsystems.

Recently, theoretical research has been undertaken to assess the ultimate nonlinear capacity of optical fibres [99], where both high level modulation formats and intra-channel nonlinearity compensating DSP has been assumed. High-level modulation formats have been a topic of much research, resulting in spectral efficiency for DP-QAM16 in excess of 7 bit/s/Hz [100]. As higher level modulation formats have relatively lower OSNR tolerance, higher launch powers are required, resulting in greater nonlinear penalties. Although research has focused on the study into both the techniques for high-order modulation and nonlinearity compensating DSP, little work has been done into the intersection of these two areas: the comparative benefits achievable with nonlinearity compensation when the order of modulation is increased.

For practical transmission systems, the dominant nonlinear impairment is due to the Kerr effect [40], which may be further subdivided into inter- and intra-channel nonlinear effects. While the mitigation of inter-channel fibre nonlinearities remains

an active research topic [101], this chapter describes the study of a digital coherent receiver and associated algorithms for the intra-channel nonlinearity compensation. Although inter-channel nonlinearities are often dominant in WDM transmission systems, the principles of mitigating them remains the same [102], and the results of this study may be applied to such systems assuming that inter-channel information sharing in the receiver is possible. To maximize the efficacy of the nonlinear DSP, a phase and polarization diverse digital coherent receiver is employed, allowing the full optical field within the receiver bandwidth to be reconstructed in the digital domain. This allows us to exploit our knowledge of the physical nature of the optical channel, and design our DSP accordingly.

This investigation is performed by a series of single-channel experiments and simulations to determine both the possible benefits and the necessary spatial resolution when using the digital backpropagation algorithm to mitigate intra-channel fibre nonlinearity. We compare the effects of nonlinearity compensation on two widely investigated high-level modulation formats: dual-polarization quadrature phase shift keying (DP-QPSK), which yields 4 bit/symbol; and dual-polarization 16-state quadrature amplitude modulation (DP-QAM16), which yields 8 bit/symbol. Both modulation formats are investigated at 10.7 GBd, such that the benefits of nonlinearity compensation may be compared with a doubling of modulation density, while the mitigating effects of dispersion and the signal bandwidth remain the same.

3.3 Experimental Transmission Setup

To characterise the functionality of our digital coherent receiver with nonlinearity compensation, we performed a set of transmission experiments to examine the effects of linear and nonlinear impairments. A particular focus was to investigate the effectiveness of nonlinear compensation techniques for DP-QPSK and DP-QAM16 modulation formats and the impact of varying DSP complexity on the transmission performance. The optical signals were transmitted multiple times through a single-span recirculating fibre loop, followed by coherent detection, digitization and offline digital signal processing, as shown in Figure 3.1. The loop consisted of 80.2 km SMF fibre with an overall chromatic dispersion of 1347 ps/nm and loss of 15.4 dB. The experimental procedure was similar to that described in [26], [103].

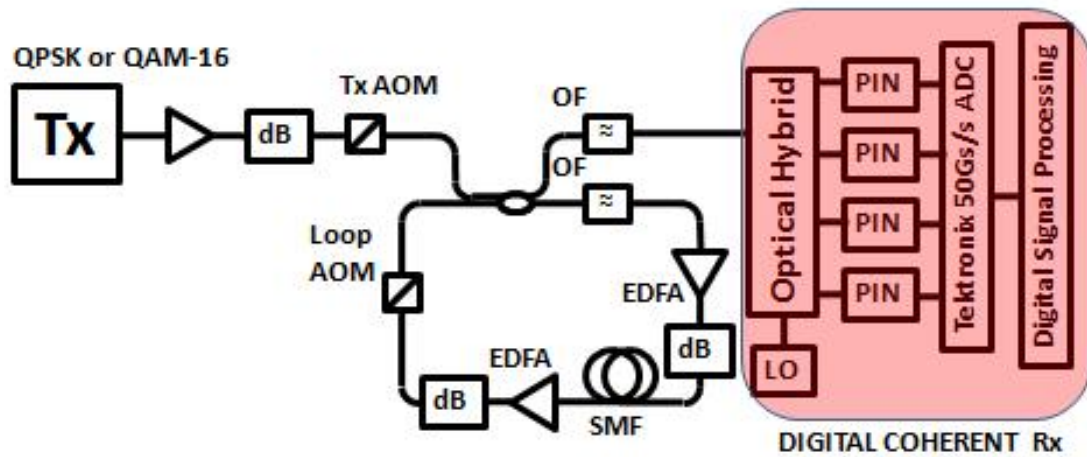


Figure 3.1 - Recirculating loop setup used for transmission experiments, with optical front end of the phase and polarization diverse digital coherent receiver.

The polarization-multiplexed QPSK signal was generated using an I-Q modulator, which was driven over $2V_{\pi}$ with respect to the minimum bias point of its transfer function. For the data two decorrelated 2^{12} PRBS sequences were used from the output of the pulse pattern generator (PPG), which were subsequently amplified to $7V_{p-p}$ ($2V_{\pi}$) to separately drive the I and Q arms of the modulator. The transmitter DFB laser linewidth, wavelength and output optical power were 1 MHz, 1554 nm and 8 dBm, respectively. To emulate polarization multiplexing we used a passive delay-line fibre interferometer, where two single polarization QPSK signals were again decorrelated, time and amplitude aligned and finally recombined via a polarization beam splitter (PBS) as shown in Figure 3.2.

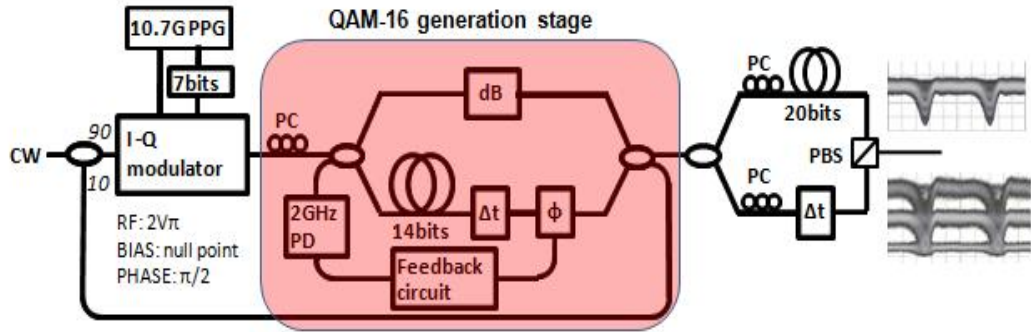


Figure 3.2 - Transmitter structure for DP-QPSK, with optional QAM16 stage highlighted. Inset: optical eye-diagrams at the output of the transmitter for DP-QPSK (top), and DP-QAM16 (bottom).

To synthesise a DP-QAM16 signal, we employed a recently developed method based on the interferometric optical processing of a QPSK signal, developed by S. Makovejs [60]. To aid carrier phase estimation, an external cavity laser (ECL) with a linewidth of 100 kHz was used in the QAM16 transmitter. The initial QPSK signal was launched into a phase-stabilised fibre interferometer, where the two signals are decorrelated, time-aligned and attenuated with respect to each other by 6 dB (highlighted, Figure 3.2). The phase between two arms was set to 90° and maintained utilizing a feedback circuit. For the feedback circuit we used a ditherless bias control circuit; alternatively, a circuit design described in [105] may be used. Even though this method cannot be used to independently modulate different streams of data, this can be used to investigate transmission performance of DP-QAM16 signals. In addition, this generation method allows suppression of the transfer of noise between the electrical and optical domains in the transmitter, owing to the nonlinear transfer function of the modulator. Polarization multiplexing emulation was performed as in the DP-QPSK case.

3.4 Simulation of Nonlinear Equaliser Performance

To verify experimental results, a 2^{15} long symbol sequence was simulated for 10.7 GBd DP-QPSK and DP-QAM16 transmission by C. Behrens. The results of these simulations were originally published with the experimental results in [12], and are therefore replicated in this chapter for the sake of completeness. Laser phase noise was modelled as a Wiener process, leading to a transmitter laser linewidth of 1 MHz for DP-QPSK and 100 kHz for DP-QAM16, while the influence of relative intensity noise (RIN) was neglected throughout the simulations. An electrical 5th order Bessel-filter with a 3 dB bandwidth of 26 GHz was used to emulate the limited transmitter bandwidth.

α [dB/km]	0.19
D [ps/km/nm]	16.87
γ [1/W/km]	1.2
PMD coefficient [ps/km ^{0.5}]	0.1
Span length [km]	80.2
Number of spans	97/20
Optical filter BW [GHz]	100
EDFA noise figure [dB]	4.5

Table 4 - Fibre and Link Parameters

To ensure accurate simulation of the experimental setup, the transmission-link in Figure 3.1 has been modelled in as much detail as possible, while making the following assumptions. Each AOM is assumed to introduce a loss of 3 dB, while EDFAs are operated in saturation to give a fixed output power of 17 dBm, and add noise power to the signal corresponding to a noise figure of 4.5 dB. Further attenuation is then applied to attain the desired launch power into the fibre. We simulated a span length of 80.2 km, modelling propagation inside the transmission fibre with a symmetrical split-step Fourier method covering linear effects, the Kerr effect, polarization mode dispersion and nonlinear polarization scattering using a nonlinear step-size of 0.1 km. The fibre parameters that were used are detailed in

Table 4 The optical loop filter was modelled as a 2nd order Gaussian filter with a 3 dB bandwidth of 100 GHz.

After transmission, the signal was detected with a single-ended coherent receiver assuming a local oscillator (LO) to signal ratio of 24 dB and an LO linewidth of 100 kHz. Limited receiver bandwidth was largely determined by the P-I-N photodiodes, which are modelled with 5th order Bessel filters using a 3 dB bandwidth of 7 GHz. Receiver-side ADCs introduce additional quantization noise and are modelled as having an effective resolution of 4 bits.

The residual implementation penalty was assumed to stem mainly from electrical noise in transmitter and receiver and was modelled by adding additional electrical noise at the receiver to give back-to-back performance similar to the measured receiver sensitivity.

3.5 DP-QPSK Transmission Results

To examine the variation of system performance with various implementations of digital backpropagation, we performed experiments with DP-QPSK near to the maximum reach without nonlinearity compensating DSP. Algorithm performance was examined over 97 spans (7780 km). As this distance is close to maximum reach, both nonlinear effects and the possible benefits of nonlinearity compensation are more significant than for shorter distances. A symbol rate of 10.7 GBd was chosen to exploit the full receiver bandwidth when using T/2 sampling for processing. The digital backpropagation algorithm was then investigated in terms of both the nonlinear step size and the use of both Wiener and Wiener-Hammerstein models.

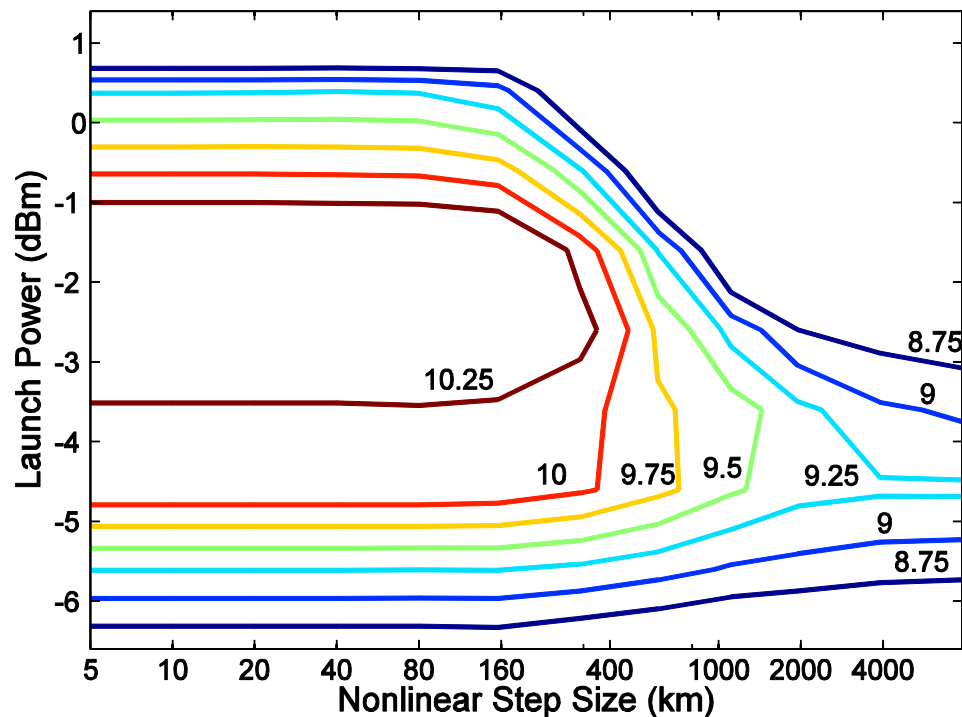


Figure 3.3 - Contour plot of experimentally determined Q-factor in dB against launch power and nonlinear step-size for Wiener cascade compensation of 97 spans transmission DP-QPSK at 10.7 GBd. Nonlinear step-size of a single span lies at 80.2 km.

The performance of the Wiener cascade model backpropagation was experimentally characterised for DP-QPSK and the results shown in Figure 3.3. Q-factor in dB is plotted as a contour graph against nonlinear step size in km on the horizontal axis and launch power in dBm on the vertical axis.

In Figure 3.3 we observe that for lower powers, performance is limited by the accumulated optical noise. For launch powers of below -5 dBm there is an insignificant improvement in performance for either modulation format with any step size. As launch power is increased, we note that an improvement in performance is available for reduced step sizes up to 160 km. The optimum launch power is improved by some 2 dB, from approximately -4.5 dBm to approximately -2.5 dBm. The benefits available with decreasing nonlinear step-size become saturated at 160 km. It is noted that while the improvement in maximum Q-factor may be modest (in the region of 1.5 dB), the increase in input dynamic range (that is, the range of launch powers for which the BER is less than the FEC limit) is more dramatic: approximately 4 dB.

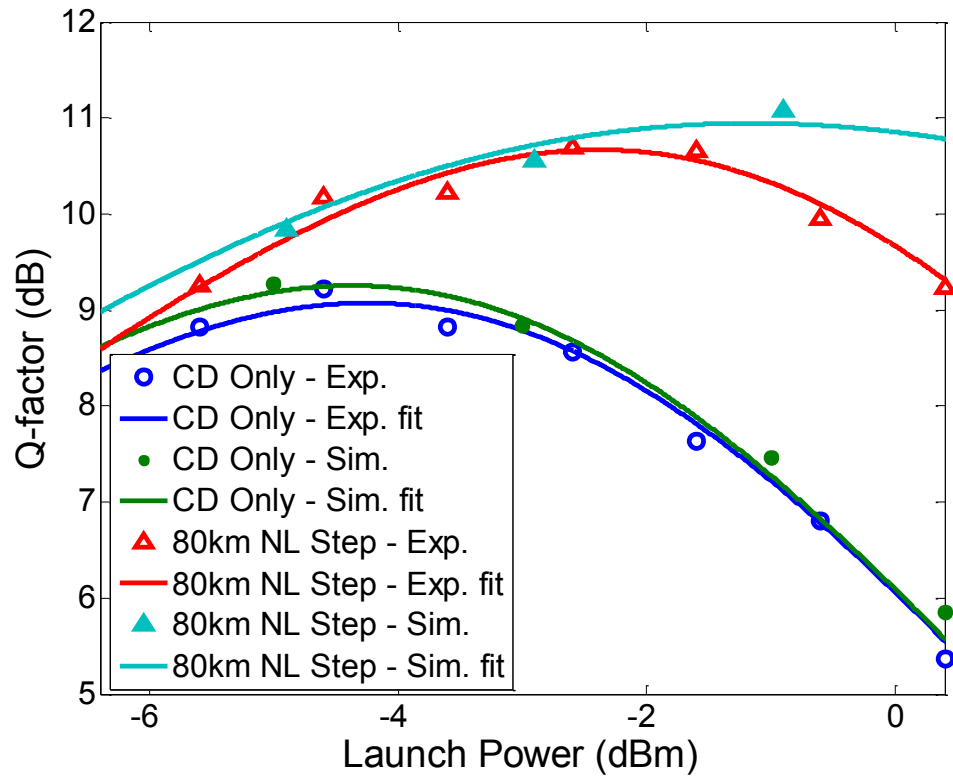


Figure 3.4 - Q-factor in dB for transmission of 10.7 GBd DP-QPSK over 97 spans. Unfilled markers denote experimental data, while solid markers correspond to simulated results, and lines denote polynomial fits. CD only and 1 step per span are demonstrated.

Figure 3.4 shows the variation of the Q-factor against launch power with and without backpropagation for a single nonlinear step per span. Experimental data is denoted by unfilled markers, while Monte-Carlo simulations (as described in the previous section) are denoted by filled markers. Simulation parameters were determined as follows: fibre chromatic dispersion, length and attenuation were measured while optical filter bandwidths and fibre PMD were taken from manufacturers specifications. The fibre nonlinear coefficient and optical amplifier noise figures were used as fitting parameters to ensure good agreement between simulated and experimental results. While the simulated results provide a good fit to the experimental results for dispersion compensation only, for short step-sizes and long transmission distances this quality of fit is much harder to attain. We believe that this is due to the unmodelled distortions in the transmitter and receiver and their interaction with the signal during backpropagation. It is noted that there is a very good agreement between the experimental and simulated results. Linear

compensation denotes compensation for chromatic dispersion only, while 1 step per span correspond to a nonlinear step-size of 80.2 km.

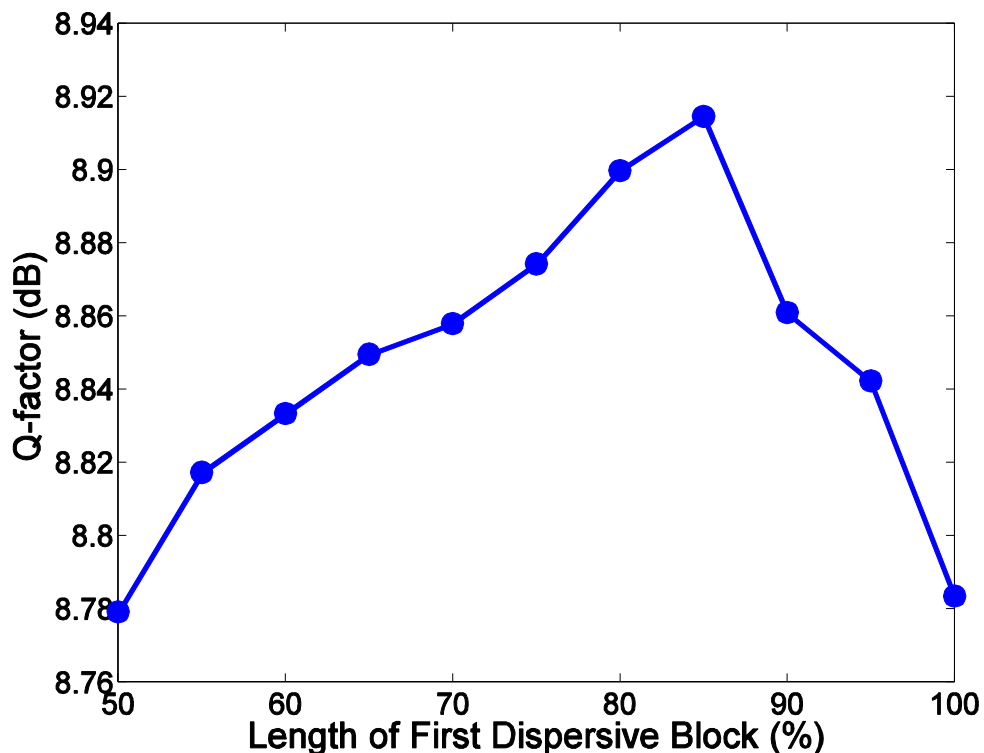


Figure 3.5 - Variation of experimental Q-factor in dB with distribution of dispersion in Wiener-Hammerstein cascade compensation of 97 spans transmission DP-QPSK at 10.7 GBd. Shown for 0.5 dBm launch power and an 80.2 km step-size.

The influence of the distribution of dispersion between the two linear blocks in the Wiener-Hammerstein cascade was then investigated. This corresponds to varying the position of the nonlinear block within the section of fibre which is approximated by each 3 block system. The optimum was found to be 85% of the dispersion in the first block and 15% in the second block. This corresponds to applying the nonlinearity at half of the effective nonlinear length of the fibre section. A graph illustrating this is shown in Figure 3.5, which shows the variation of Q-factor with the length of the first dispersive step. The figure shows an 80 km nonlinear step size for 0.5 dBm launch power. Other launch powers and nonlinear step sizes also exhibit a maximum when the split of dispersion is 85% - 15%, although the maximum improvement in Q-factor is smaller for lower launch powers and smaller nonlinear step sizes. This dispersive split was used for all subsequent uses of Wiener-Hammerstein model backpropagation. While the demonstrated maximum improvement in the optimal

case is small in Figure 3.5, this improvement becomes greater as the nonlinear step-size increases.

To effectively represent and compare the results between the different backpropagation models, we used fitted curves to quantify improvements in launch power and Q-factor with reduction in step size. A least squares fit was used to produce the polynomial approximation of Q-factor as a function of launch power given in (3.1) [106].

$$\frac{1}{Q^2} = \frac{1}{aP} + \frac{P^x}{b} \quad (3.1)$$

Using this fitting process, we were able to extrapolate the optimum launch power for different models and spatial resolutions of nonlinearity compensation. This inferred optimum launch power is useful as the experimental measurements have a granularity of 1 dB in launch power, and the change in optimum launch power may be less than this between different nonlinear step sizes. From the polynomial approximation of variation in Q-factor with launch power we are also able to infer Q-factor at the optimum launch power.

To account for the difference in complexity between the two nonlinear models and allow a more direct comparison, performance is characterised for the mean dispersive block length, which we define as (3.2):

$$\text{Mean Dispersive Block Length} = \frac{\text{Total Link Length}}{\text{Number of Dispersive Bloc}} \quad (3.2)$$

Where there are two adjacent dispersive blocks in the Wiener-Hammerstein system, we assume that they may be incorporated into a single block. Therefore, for a very high order cascade, the mean dispersive block length of the Wiener-Hammerstein system will approach that of the Wiener system. Conversely, for a first order nonlinear model, the Wiener model will have a mean dispersive block length which is twice that of the Wiener-Hammerstein model.

The characterization of DP-QPSK with varying mean dispersive block length is given below in Figure 3.6 and Figure 3.7 for both Wiener model and Wiener-

Hammerstein model backpropagation. Here we plot improvement in Q-factor at the optimum launch power (Figure 3.6) and optimum launch power improvement (Figure 3.7) using the polynomial fitting process with the experimental measurements previously described.

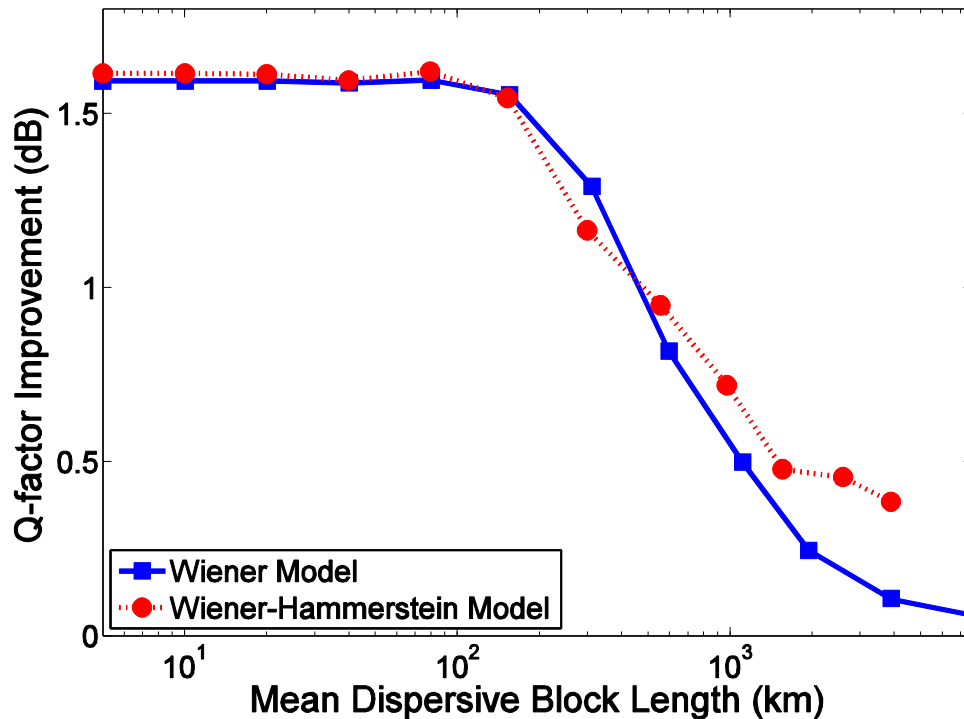


Figure 3.6 - Plot of improvement in inferred maximum Q-factor against mean dispersive block length for DP-QPSK at 10.7 GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.

It is noted that in both Figure 3.6 and Figure 3.7 both nonlinear compensation models offer similar potential benefits for a given mean dispersive block length. Maximum Q-factor is improved by approximately 1.6 dB, and optimum launch power is increased by approximately 1.9 dB for both compensation models. The improvement in both Q-factor and launch power saturates for a mean dispersive block length of approximately 160 km, or two spans. The similarity between the models when characterised in terms of mean dispersive block length may be considered with reference to the discussion of the accuracy of the models presented in the literature review. When the step size is small, the accuracy of both models is good and there is agreement in accuracy between the two models as the error due to the commutability of the linear and nonlinear blocks is insignificant. As the

nonlinear step-size increases, the benefit in accuracy of the Wiener-Hammerstein model becomes more significant. This difference is offset, however, due to the difference in mean dispersive block length between the two models. This becomes most noticeable for a nonlinear step-size corresponding to the link length L : for this case the mean dispersive block length is L for the Wiener cascade, but $L/2$ for the Wiener-Hammerstein cascade.

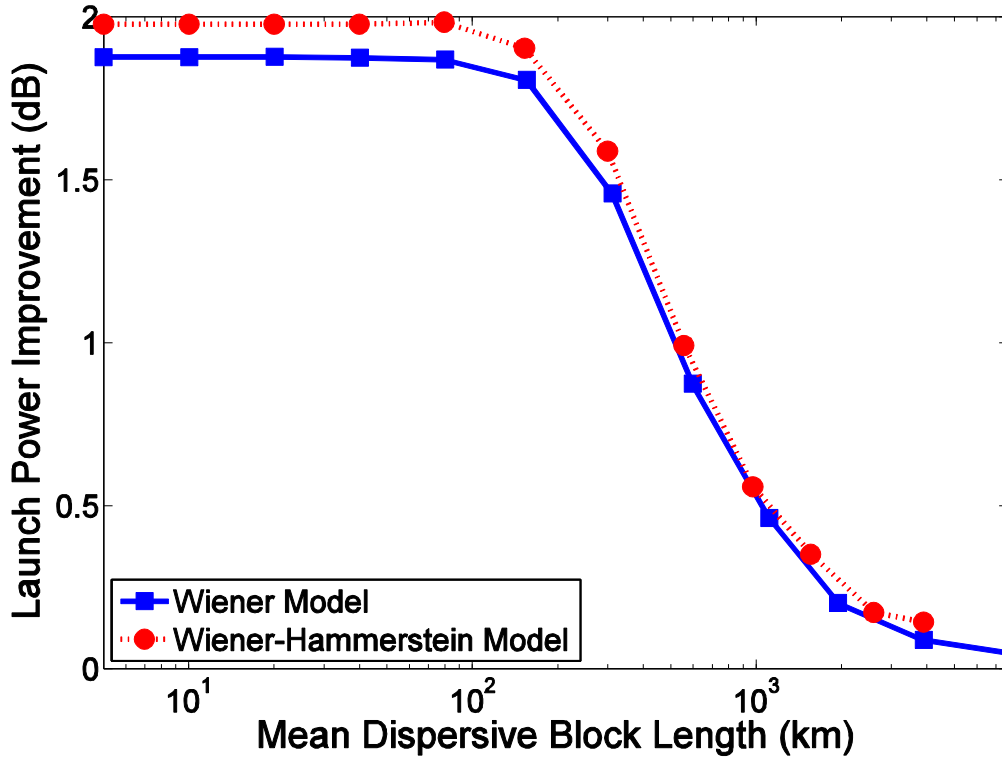


Figure 3.7 - Plot of improvement in inferred optimum launch power against mean dispersive block length for DP-QPSK at 10.7 GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.

3.6 DP-QAM16 Transmission Results

We then investigated the performance benefits available when using the digital backpropagation algorithm for the compensation of fibre nonlinearity in a DP-QAM16 system over 20 spans (1600 km). Again, this transmission distance was chosen to be close to maximum reach with linear DSP, in order that the benefits available from nonlinear equalisation are more noticeable. To enable comparison of the modulation formats, we performed the comparison at the same symbol rate

(10.7 GBd), thus ensuring that the nonlinearity-mitigating effect of chromatic dispersion is identical for both formats.

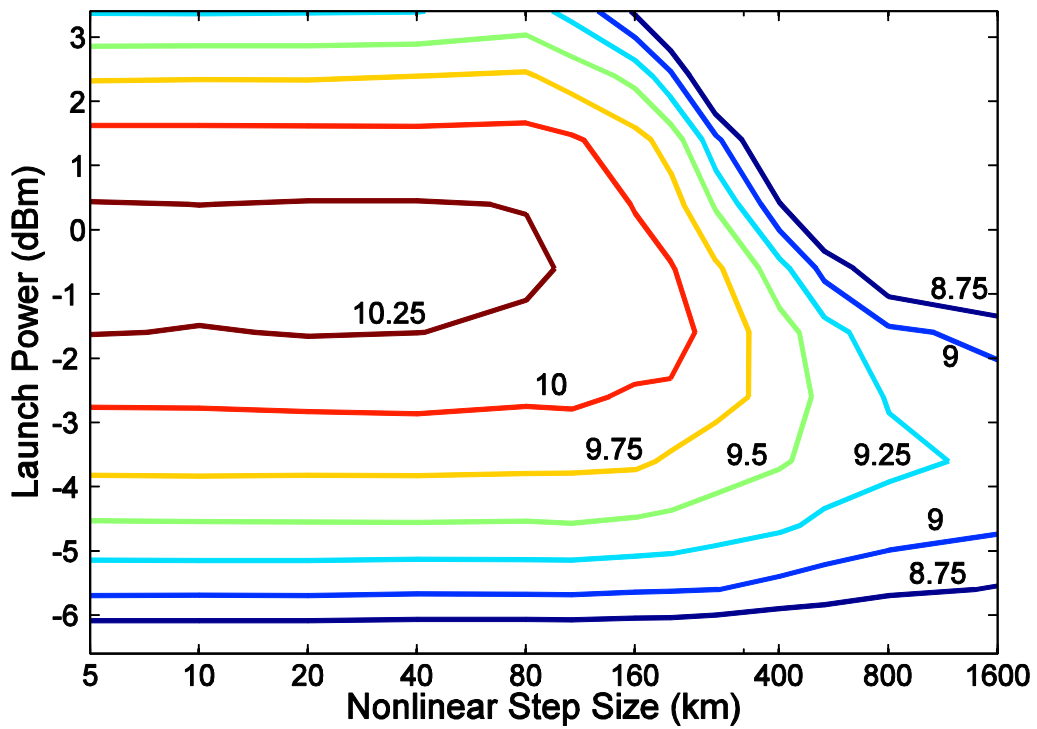


Figure 3.8 - Contour plot of experimentally determined Q-factor in dB against launch power and nonlinear step-size for Wiener cascade compensation of 20 spans transmission DP-QAM16 at 10.7 GBd. Nonlinear step-size of a single span lies at 80 km.

As in the previous section, we applied Wiener model backpropagation applied to DP-QAM16 transmission varying nonlinear step-size and launch power. These results are presented in a contour plot in Figure 3.8. Again we see that a significant improvement in performance is available with nonlinear backpropagation. In this case, the improvement in performance which is obtained by reducing the nonlinear step-size is saturated for a step-size of approximately 80 km, compared to 160 km for DP-QPSK. Similarly in the large step-size region, a smaller step-size than DP-QPSK is required for DP-QAM16.

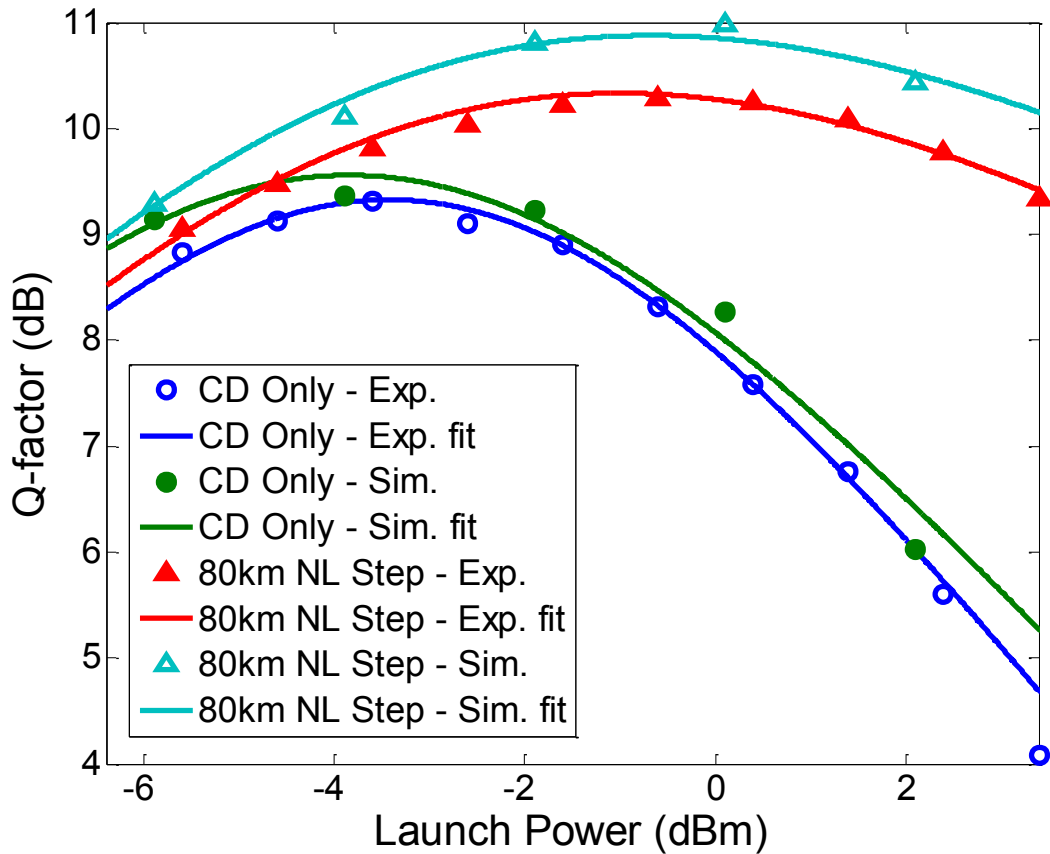


Figure 3.9 - Q-factor in dB for transmission of 10.7 GBd DP-QAM16 over 20 spans. Unfilled markers denote experimental data, while solid markers correspond to simulated results, and lines denote polynomial fits. CD only and 80 km nonlinear step-size are demonstrated.

Again, simulations were performed in Matlab to determine the agreement of the experimental data with theoretical models. Figure 3.9 shows the variation of the Q-factor against launch power with and without nonlinear backpropagation. Experimental data is denoted by unfilled markers, while Monte-Carlo simulations (as described in the previous section) are denoted by filled markers. It is noted that there is a very good agreement between the experimental and simulated results. CD only denotes compensation for chromatic dispersion only, while 1 step per span corresponds to a nonlinear step size of 80 km.

Again, we employed the polynomial fit described in equation (3.2) to the experimentally obtained data to ascertain the optimum launch power and Q-factor for each nonlinear step-size. These results were then plotted for both Wiener model backpropagation and Wiener-Hammerstein model backpropagation. These results are presented in Figure 3.10 and Figure 3.11.

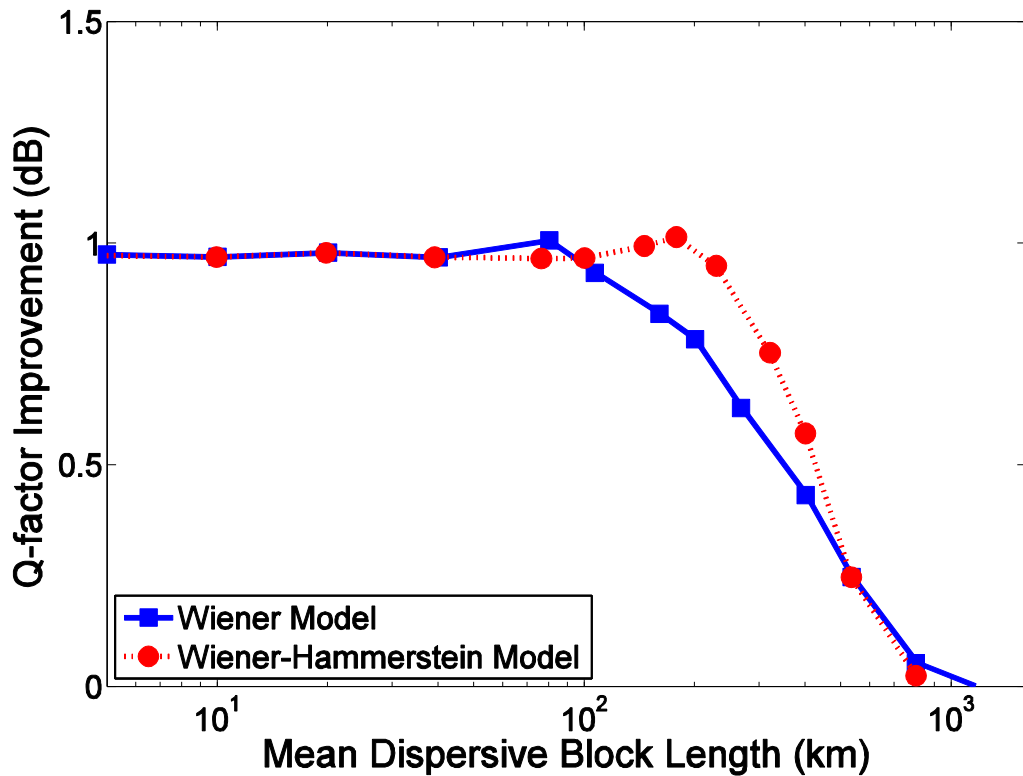


Figure 3.10 - Plot of improvement in inferred maximum Q-factor against mean dispersive block length for DP-QAM16 at 10.7 GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.

We note from Figure 3.10 and Figure 3.11 that similarly to DP-QPSK, both nonlinear compensation models offer similar potential benefits for a given mean dispersive block length. The maximum available benefit in Q-factor is approximately 1 dB, and optimum launch power is increased by approximately 2.5 dB for both compensation models. The improvement in both Q-factor and launch power saturates for a mean dispersive block length of approximately 80 km, or one span.

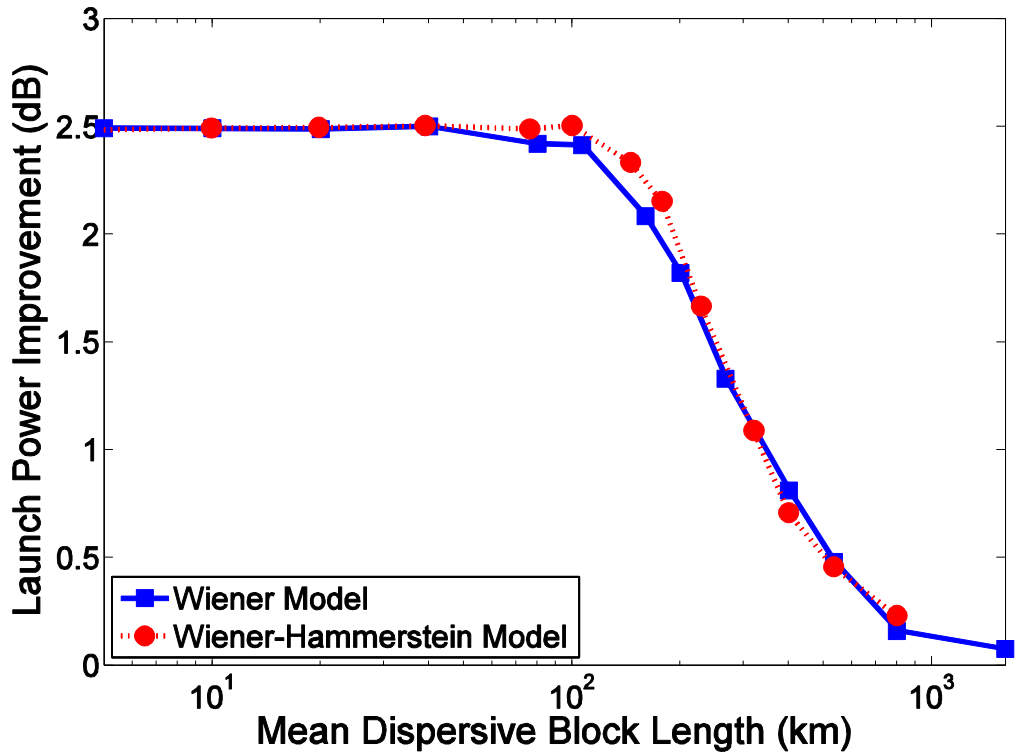


Figure 3.11 - Plot of improvement in inferred optimum launch power against mean dispersive block length for DP-QAM16 at 10.7GBd using Wiener and Wiener-Hammerstein model nonlinearity compensation.

3.7 Comparison of Performance for DP-QPSK and DP-QAM16

By noting that the available benefit from Wiener-Hammerstein model backpropagation is almost identical to Wiener model backpropagation for both DP-QPSK and DP-QAM16 systems, we may justify comparing between modulation formats using only one model. For this reason we will compare performance of backpropagation using the most prevalent model in the literature, the Wiener model. As a metric for performance improvement, we will discuss the increase in optimum launch power, which is approximately proportional to the launch power margin.

The comparison between DP-QPSK and DP-QAM16 in terms of the number of nonlinear block required is presented in Figure 3.12, while a comparison by mean dispersive block length is presented in Figure 3.13.

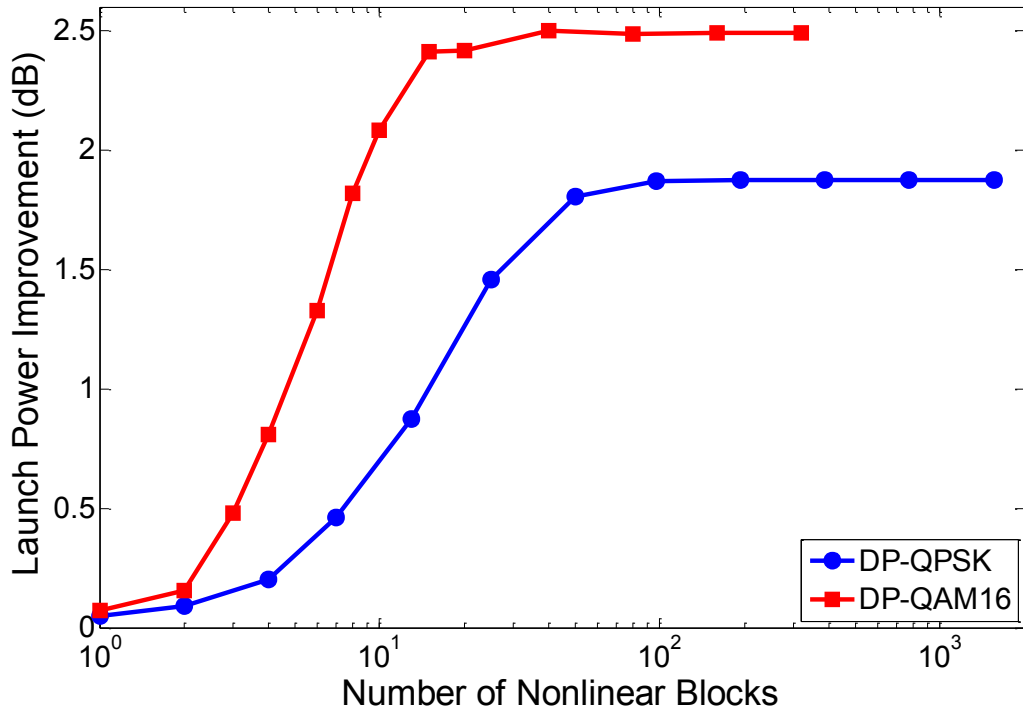


Figure 3.12 - Plot of improvement in inferred optimum launch power against number of nonlinear blocks for Wiener model nonlinearity compensation of DP-QPSK and DP-QAM16.

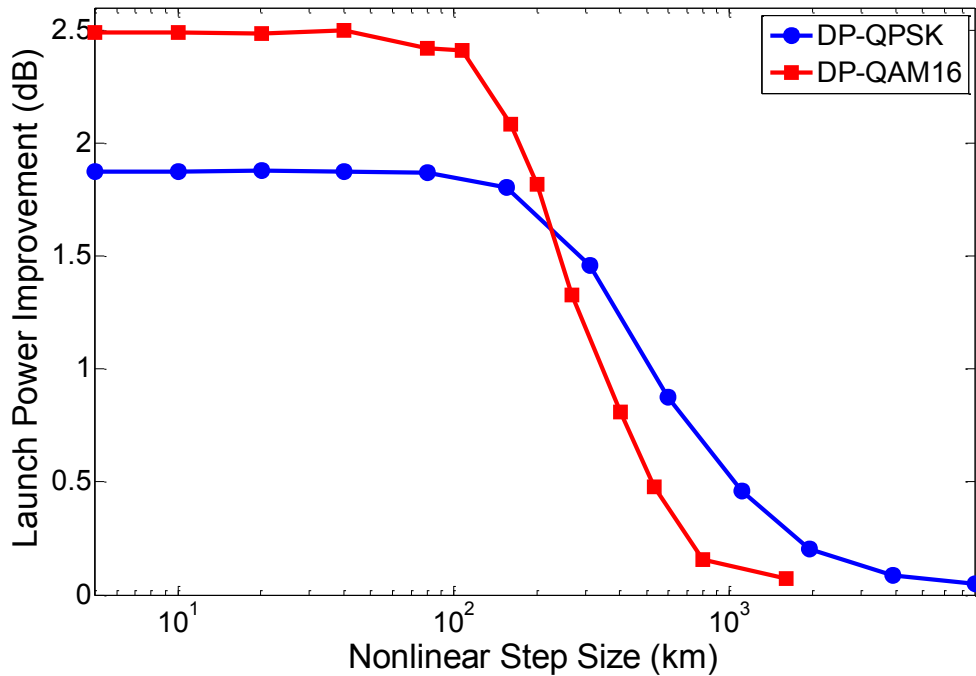


Figure 3.13 - Plot of improvement in inferred optimum launch power against nonlinear step size for Wiener model nonlinearity compensation of DP-QPSK and DP-QAM16.

We note from Figure 3.12 that for a 1.5 dB increase in optimum launch power, that DP-QPSK over 97 spans (7780 km) requires approximately 25 nonlinear blocks, while DP-QAM16 requires approximately 7 over 20 spans (1600 km). This corresponds to a slightly smaller nonlinear step-size for DP-QAM16 (230 km compared to 310 km) as may be seen in Figure 3.13, though this difference is dwarfed by the reduction of nonlinear blocks due to the reduced transmission distance. This is also in agreement with our earlier observation that DP-QAM16 requires a smaller nonlinear step size than DP-QPSK to gain the maximum available benefits from backpropagation. We note that in future coherent systems which employ highly dense modulation and relatively short transmission distances, nonlinearity compensating DSP is likely to be more attractive than for current systems, which are largely DP-QPSK.

3.8 Estimated Implementation Complexity for Nonlinear Backpropagation

The computational complexity of digital backpropagation may be approximated in the following manner: the number of complex multipliers in the dispersion compensating (linear) elements will dominate the implementation complexity of backpropagation. Therefore, the majority of the increase in complexity when reducing the nonlinear step-size is due to the increased number of operations resulting from the number of FFTs required (if we assume dispersion compensation is to be performed in the frequency domain). We may therefore use the approximation given by (3.3), which describes the number of complex multiplications per output sample when the radix 2 FFT is used [107]:

$$N_{CMAC} \propto \frac{n \log_2(n)}{s} \quad (3.3)$$

Where N_{CMAC} is the number of complex multiplications per output sample, n is the FFT length, and s is the overlap length used in the filtering operation. When applied to the step-size analysis in the results in the previous sections (initially published in [12]), and filtering at two samples per symbol, we note that in single channel systems, the benefit in launch power is approximately linearly related to the relative complexity increase in the region prior to improvement saturation. This effect is illustrated by Figure 3.14, where we assumed an FFT size of 2^{12} and an overlap of 25%.

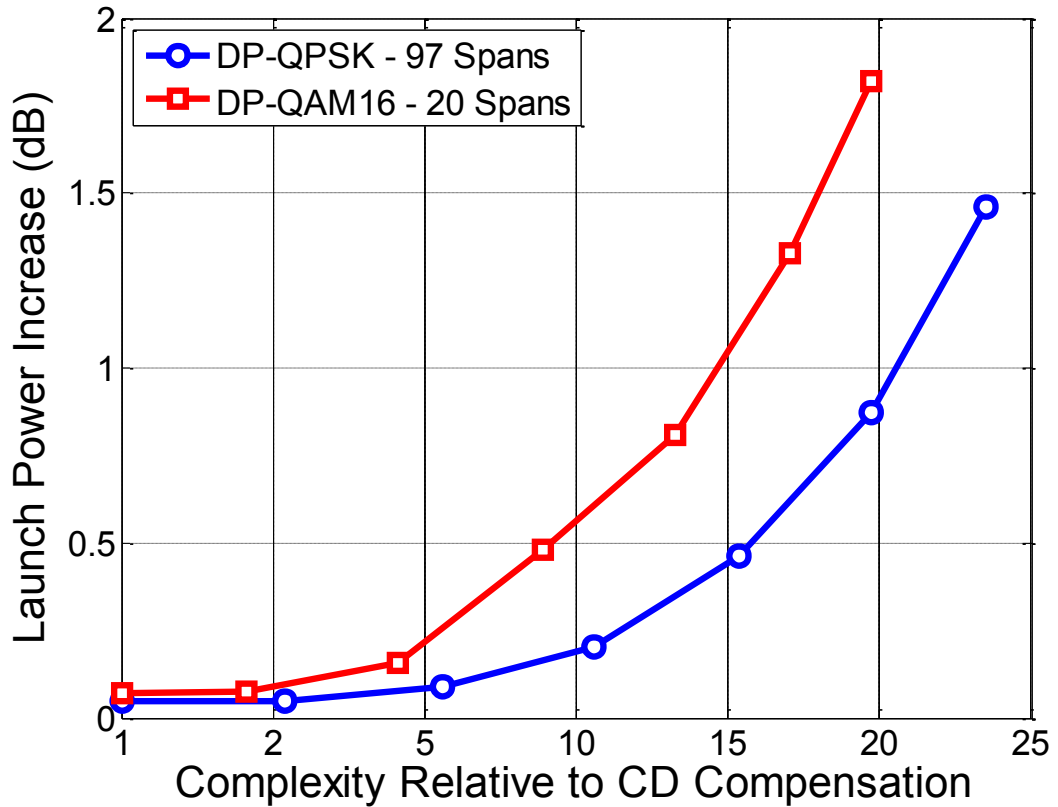


Figure 3.14 - Relative increase in complexity for backpropagation of 10.7 GBd DP-QPSK and DP-QAM16

We note from Figure 3.14 that that in the case of 10.7 GBd DP-QAM16, an increase of a factor of 15 in the number of complex multiplications is required for a 1 dB increase in optimum launch power. Similarly, in the case of 10.7 GBd DP-QPSK, an increase of a factor of 20 in the number of complex multiplications is required for a 1 dB increase in optimum launch power. It is also noted that for a single nonlinear step, there is a small improvement in performance over CD compensation only for a negligible increase in complexity. Since the static CD compensating filter is known to contribute a significant complexity to the receiver ASIC, even at 10.7 GBd [108], this increase in complexity is highly significant to the feasibility of backpropagation for implementation in commercial products. Since we may consider that single channel operation is an upper bound to the improvement in performance which may be achieved with a WDM system, we note that the improvement in performance is likely to be minimal while highly costly in terms of computational complexity (and therefore ASIC design cost and power consumption).

3.9 Summary

In this chapter, we have investigated the performance of a coherent receiver with nonlinearity-compensating DSP and have shown that it can be successfully used to mitigate intra-channel nonlinearities in both DP-QPSK and DP-QAM16 over distances of 7780 km and 1600 km respectively. The impact of the key receiver DSP parameter, namely the nonlinear step size was investigated. It was shown that significant improvements in performance may be achieved with resolution significantly coarser than a single span. While performance in this *long-step* region may be improved with the use of a three block Wiener-Hammerstein model rather than the more commonly used Wiener model, the increased computational effort this model requires offsets any benefit when performance is examined in terms of the mean dispersive block length. For the examined receiver bandwidth and symbol rate, the benefit of nonlinearity compensation saturated for a nonlinear step size of 160 km for DP-QPSK and 80 km for DP-QAM16. Additionally, nonlinear backpropagation appears to offer a greater benefit for DP-QAM16, which may be attributed to this format's higher susceptibility to fiber nonlinear effects. This leads us to infer that nonlinearity compensation of this kind is considerably more attractive for modulation formats which are highly spectrally efficient, and transmitted over short links, where the reduced memory due to dispersion and the increased benefit available combine to produce greater benefits from fewer nonlinear steps. We then compared the computational complexity of nonlinear backpropagation to that of chromatic dispersion compensation only. An approximately exponential relationship was found between complexity in terms of the required number of complex multipliers and performance in Q-factor for both modulation formats. A 1 dBQ improvement in performance requires an increase in complexity of approximately a factor of 10 for DP-QAM16 and a factor of approximately 15 for DP-QPSK. This result indicates that even in the single channel case, the available improvement in performance when using nonlinear backpropagation is limited, while the computational cost is high.

While the benefits of nonlinearity compensation may be considerable in some cases, the computational cost of implementation makes these algorithms prohibitive at the present time. While digital nonlinearity compensation is at best a distant possibility, other possibilities exist to provide better performance than conventional coherent

DP-QPSK systems at 40G and 100G line rates. In the following sections, we will investigate the processing and performance of a modulation format which has optimal power constrained performance in four dimensions: polarization-switched QPSK.

In the next chapter we will examine the digital post-processing of a modulation format known as polarization-switched QPSK (PS-QPSK). This modulation format has superior noise tolerance to DP-QPSK, and will therefore offer performance benefits in both the linear and nonlinear regions without the need for the computationally intensive algorithms required for digital backpropagation.

4 Polarization Switched QPSK: Theory and Digital Equalisation

4.1 Abstract

Coherent detection in combination with digital signal processing has recently enabled significant progress in the capacity of optical communications systems. This improvement has enabled detection of optimum power-constrained modulation formats for optical signals in four dimensions. The increased noise tolerance of these modulation formats results in superior transmission performance which may be considered as preferable alternative to digital nonlinearity compensation, as the complexity of implementation is much lower. In this chapter, we investigate digital post-processing of one such modulation format: polarization-switched quadrature phase shift keying (PS-QPSK). Coding schemes, equalisation and WDM sensitivity are analysed, and a novel equalisation algorithm investigated. The proposed algorithm, which includes both blind initialisation and adaptation of the equaliser, is found to be insensitive to the input polarization state and demonstrates highly robust convergence in the presence of PDL, DGD and polarization rotation.

4.2 Introduction

Polarization and phase diverse coherent detection with digital signal processing has become an essential technique for mitigating fibre transmission impairments and therefore increasing capacity [109]. The basis of polarization and phase diverse coherent detection is that the in-phase and quadrature components of the two orthogonal polarizations are detected, corresponding to all four dimensions of the incoming optical field [104]. As all four dimensions of the field are detected, transmission impairments such as chromatic dispersion (CD) and polarization mode dispersion (PMD) may be compensated [10]. Recently, much research has been performed into the compensation of self phase modulation (SPM) in order to gain improvements in performance of 1 or 2 dB at great computational cost [12], [63], [110].

While the detection of all four dimensions of the incoming optical field has enabled mitigation of transmission impairments, it has also enabled the use of high-level modulation formats such as quadrature phase shift keying (QPSK) [18] and 16-state quadrature amplitude modulation (QAM16) [100]. These modulation formats are most commonly transmitted on two orthogonal linear polarizations, doubling the achievable spectral efficiency. This set of modulation formats are often denoted as dual-polarization (DP-) or polarization (division) multiplexed (PDM-; PM-; or PolMux-).

Recently, research has been performed into determining the optimum modulation format in four dimensions, given that we have the ability to detect and digitally process all four dimensions of the transmitted optical field [111]. Previous proposals have focused on optimal constellations for the power constrained case [34], with some research being performed into using the extra capacity afforded by using an optimal 24-state constellation as coding overhead [112]. More recently, research into the performance of PS-QPSK in transmission has entered the literature. While this may seem to run against the trend for higher levels of modulation and more dense constellations, it may be noted that in industry there is still a demand for highly robust transmission at the expense of spectral efficiency for ultra long-haul applications [113], [114]. Despite the recent interest in PS-QPSK modulation, DSP algorithms specifically designed for PS-QPSK were published only while the

algorithm presented in this chapter [13] was in peer review [14] and shortly after publication [115]. Research prior to this focusing on transmission performance of PS-QPSK over uncompensated links [116] used an equaliser utilising a training sequence. The work presented in [117] examined the performance of PS-QPSK over dispersion managed links, did not discuss equalisation. The work presented in these papers [116], [117] indicates that PS-QPSK modulation offers a significant advantage over DP-QPSK in transmission at 112 Gb/s.

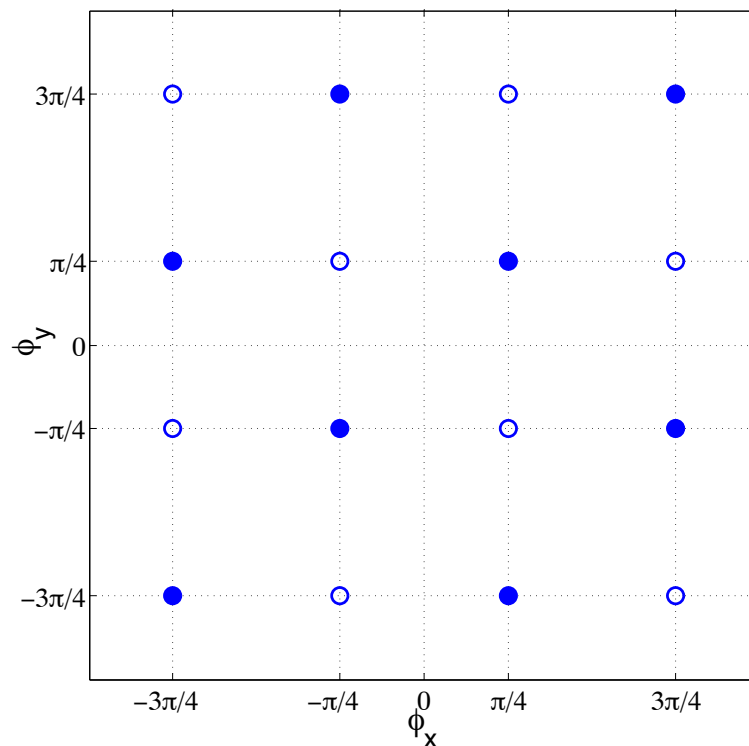


Figure 4.1 - Constellation diagrams for DP-QPSK and PS-QPSK in the phase space.

Phase on the x polarization (ϕ_x) is plotted against phase on the y polarization (ϕ_y) (both in radians). Solid markers belong to both formats while hollow markers belong to PS-QPSK only.

PS-QPSK may be visualised as a subset of DP-QPSK under a polarization rotation. This is illustrated in Figure 4.1 where the phase of each polarization of both formats is plotted (PS-QPSK is rotated relative to the representation given elsewhere in this thesis for clarity).

In this chapter we will examine the coding and digital post-processing of PS-QPSK, which is the optimal 8-state constellation. PS-QPSK offers the greatest benefit in

performance of any of the novel four-dimensional modulation formats proposed thus far, and also offers unique challenges and benefits in the design of digital subsystems.

4.3 PS-QPSK Modulation and Coding Performance

The PS-QPSK constellation may be considered as the 4-dimensional regular polytope (polychoron) known as the tesseract, or 8-cell. This may be more easily visualised as a QPSK constellation which is transmitted on one of two orthogonal polarizations [111]. We may also consider the constellation as the following set of vectors in \mathbb{C}^2 space (4.1):

$$S_{PS-QPSK} \in \left\{ \begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \pm j \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm j \end{pmatrix} \right\} \quad (4.1)$$

It has been demonstrated that this modulation format has superior symbol error rate (SER) performance compared with conventional modulation formats used in current communication systems. This previous analysis bears repetition, as it relates the fundamental advantage of using a modulation format which is optimal under constrained power in four dimensions. An informative metric with which to compare performance of commonly used modulation formats with PS-QPSK is the variation of SER with SNR per transmitted bit. This will relate achievable performance (SER) to the achievable SNR given a desired gross bit rate (E_b/N_0). SER for PS-QPSK may be given by (4.2), while SER for DP-QPSK is given by (4.3).

$$SER_{PS-QPSK} = 1 - \frac{1}{\sqrt{\pi}} \int_0^{\infty} (1 - \operatorname{erfc}(x))^3 e^{-\left(x - \sqrt{\frac{3E_b}{N_0}}\right)^2} dx \quad (4.2)$$

$$SER_{DP-QPSK} = 1 - \left[1 - \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{E_b}{N_0}} \right) \right]^4 \quad (4.1)$$

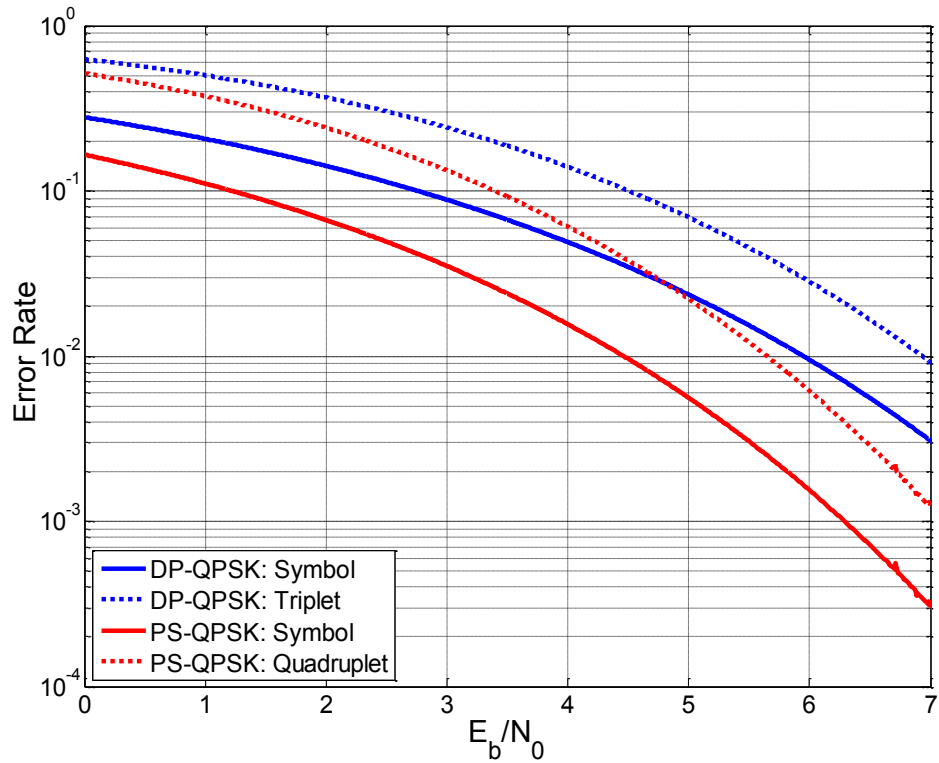


Figure 4.2 - Ideal receiver sensitivity in symbol error rate against E_b/N_0 for DP-QPSK and PS-QPSK. For equivalent information carrying capacity, error rate of triplets of DP-QPSK is compared with error rate of quadruplets of PS-QPSK.

We note immediately from Figure 4.2 that for identical modulation rates, PS-QPSK has receiver sensitivity significantly better than DP-QPSK at high symbol error rates, increasing to an asymptotic improvement of 1.76 dB. This is despite the fact that PS-QPSK carries 3 bits of information per symbol while DP-QPSK carries 4 bits of information per symbol, necessitating an increase in symbol rate by a factor of 4/3 for an equal bit rate. To compare symbol error rates with equivalent information, we have also compared symbol triplet errors in DP-QPSK to symbol quadruplet errors in PS-QPSK, yielding in each case a ‘super-symbol’ which contains 12 bits, transmitted at an equal rate. We found that PS-QPSK is superior over the domain examined, giving an improvement of 1.3 dB at a ‘super-symbol’ error rate of 10^{-2} . Comparing modulation formats in this manner indicates that PS-QPSK is worthy of further investigation. The ultimate performance metric of any digital communications system is bit error rate (BER), and to examine this facet of PS-QPSK performance, we must first provide a map of bits to symbols.

Gray codes are a class of codes for which the Hamming distance between two points

which are adjacent in the signal space is 1. These codes are extensively used in digital communications in order to minimise the number of bit errors per symbol error. An interesting feature of PS-QPSK is that there exists no Gray code, as previously discussed in [34]. The Euclidean distance between any two points as described in (4.1) may be zero (the points are the same); 2 (polarization remains the same, while phase changes by π); or $\sqrt{2}$ (any other pair of points).

Karlsson and Agrell proposed the following intuitive scheme, illustrated in Figure 4.3 [34]. As each point in the PS-QPSK constellation has 6 nearest neighbours, the least-bad coding scheme is therefore to assign the 3 bit complement (which has maximum Hamming distance) to the 1 non-nearest neighbour (maximum Euclidean distance). As this mapping scheme has previously been proposed, we will refer to this scheme hereafter as ‘conventional’ mapping. Note that in addition to the expected four QPSK constellation points on each polarization, there is a fifth point at the origin. This point corresponds to the set of symbols for which the QPSK symbol exists on the orthogonal polarization.

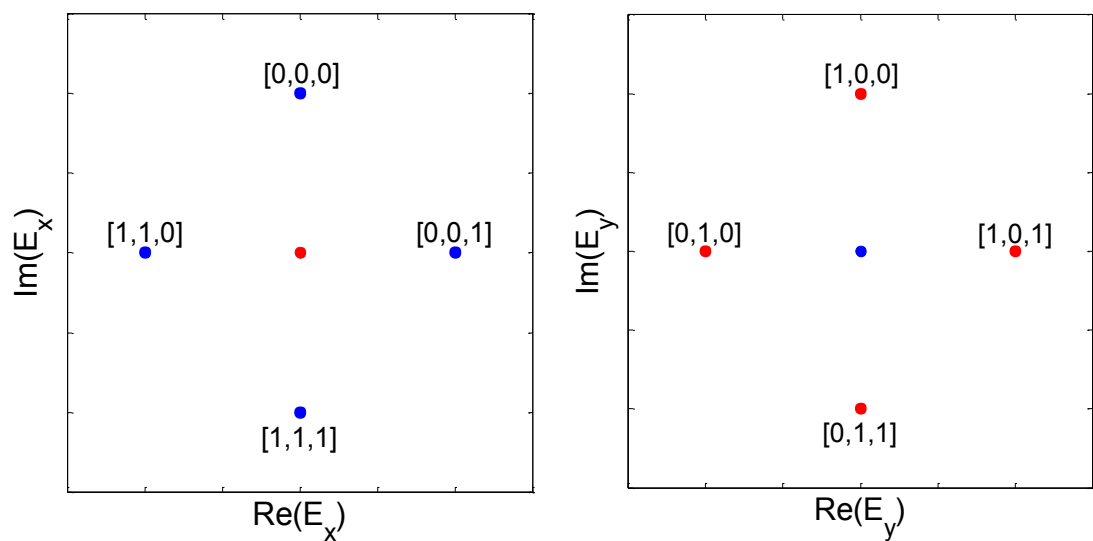


Figure 4.3 - 'Conventional' bit mapping for PS-QPSK after Karlsson & Agrell.

While the mapping scheme presented in Figure 4.3 provides an intuitive basis for coding a 4-dimensional modulation format based on an information-theoretical approach, we will briefly consider a sub-optimal bit mapping in passing. The three bits in each PS-QPSK symbol may be considered as 1 bit coding for polarization (a binary state) and two bits coding for phase (a quaternary state). We may therefore

construct a mapping which uses two Gray coded bits for the phase of the QPSK constellation and 1 bit for polarization. This scheme is shown in Figure 4.4, and will be referred to as ‘*alternative*’ mapping. Although this bit mapping is inferior to conventional coding in terms of sensitivity, it provides both a convenient intermediate step to forming a scheme for differential coding, and a simplified coding scheme for which a small performance penalty must be paid.

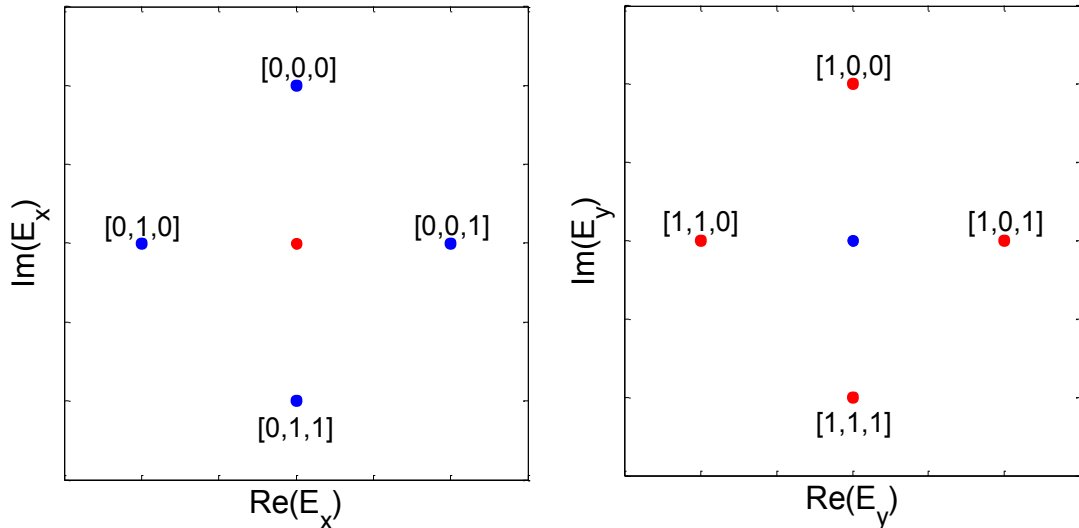


Figure 4.4 - 'Alternative' bit mapping scheme for PS-QPSK.

Differential coding is a highly effective tool, both for resolving the phase-space ambiguity present with modulation formats which possess rotational symmetries, and for reducing the effect of burst errors induced by cycle slips in the carrier phase estimation process. This scheme is widely used for M-ary PSK, and has also been proposed for more spectrally efficient QAM16 and QAM64 in the optical domain [119]. We propose a coding scheme which provides differential coding in both phase and polarization for PS-QPSK. In this coding scheme, we code two bits onto the change in phase, and one bit onto the change in polarization, providing functionality similar to that proposed in [118]. While this code is suboptimal in the same sense as the alternate bit mapping scheme described in Figure 4.4 (and also suffers from a doubling of errors seen with differentially coded M-PSK) differential coding ensures a greater robustness to the effects of phase noise and subsequent errors during carrier phase estimation. As with differentially coded DP-QPSK, the bits comprising each symbol may be differentially coded and decoded independently, though in this case, the polarization ambiguity problem is also solved. An example of the transitions

described by uncoded bits for the differential code is given in Figure 4.5. It is also worth noting that after differential coding, the map of coded bits to symbols is the same as that provided by the alternative mapping described in Figure 4.4.

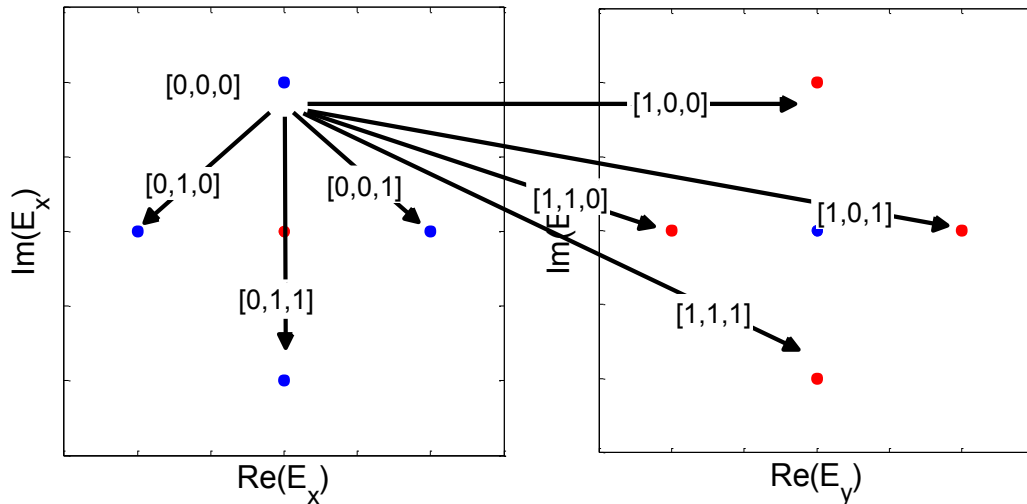


Figure 4.5 - Example of differential coding for PS-QPSK, based on 'alternate' bit mapping scheme.

To illustrate the relative performance of these coding schemes for PS-QPSK, in Figure 4.6, a BER is plotted against SNR per uncoded bit for an ideal transmitter and receiver. A benefit in required SNR of 1 dB is found between DP-QPSK and *conventional* PS-QPSK (assuming a target BER of 10^{-3}), while the *alternate* mapping yields a penalty of less than 0.05 dB. The benefit of differentially coded PS-QPSK is similarly approximately 1 dB in sensitivity over differentially coded DP-QPSK. The bit error rates for DP-QPSK were obtained from the analytical formulas provided in [35], while for PS-QPSK in each case we assumed that in the high SNR region of interest, the bit error rate is dominated by the mean Hamming distance of each of the 6 nearest neighbours [34]. In the case of the conventionally coded constellation, three of the nearest neighbours have a Hamming distance of one, while three have a Hamming distance of two. By assuming that the possibility of a non-nearest neighbour symbol error is negligible (which is a valid assumption in the high SNR regime), for a single symbol error there is a 50% chance of a 1 bit error and a 50% chance of a 2 bit error. This results in an average of 1.5 bit error per symbol error. By scaling this to the modulation density of 3 bits per symbol we find that the BER is approximately 0.5 of the SER. The same logic may be applied to the

other bit-mapping schemes, leading to the approximations given in (4.4):

$$\begin{aligned}
 BER_{Conv} &\approx \frac{1}{2} SER_{PS-QPSK} \\
 BER_{Alt} &\approx \frac{5}{9} SER_{PS-QPSK} \\
 BER_{Diff} &\approx \frac{10}{9} SER_{PS-QPSK}
 \end{aligned}
 \tag{4.4}$$

These approximations may be compared to that for DP-QPSK (4.5):

$$BER_{DP-QPSK} \approx \frac{1}{4} SER_{DP-QPSK}
 \tag{4.5}$$

While we note a benefit for PS-QPSK over DP-QPSK in the region of 1.5 dB when comparing SER at equal bit rates in Figure 4.2, this improvement is only approximately 1 dB when comparing BER.

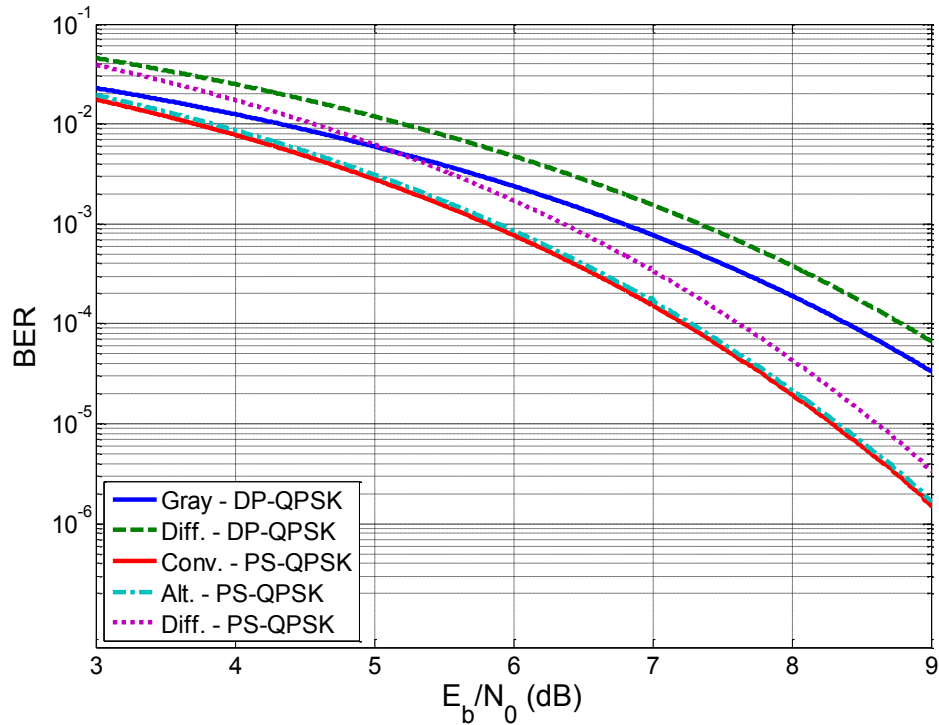


Figure 4.6 - Ideal receiver sensitivity in bit error rate (BER) against E_b/N_0 for DP-QPSK and PS-QPSK with various coding schemes and bit mappings.

Due to the slight advantage in receiver sensitivity afforded by the mapping proposed by Karlsson & Agrell (described as conventional mapping), we will use this map of bits-to-symbols in the following analysis of the polarization-switched CMA equaliser. During subsequent analysis of polarization tracking, we will also consider the impact on performance of differential coding.

4.4 The Polarization Switched CMA Equaliser

The most common equaliser in use for digital coherent communication is the constant modulus algorithm (CMA) equaliser with least mean squares (LMS) adaptation [46]. Due to the predominance of DP-QPSK as a modulation format, this algorithm is used in MIMO configuration, to separate the polarization sources and deconvolve each of these with the channel. The CMA has gained such high popularity by being robust, relatively computationally efficient and enabling the separation of the functionality of the equaliser from that of carrier phase and frequency estimation. Due to the large body of research based around the CMA and the familiarity of engineers with this algorithm, it is desirable to use this algorithm as the basis of an equaliser used for PS-QPSK.

The standard dual-polarization CMA (DP-CMA) attempts to minimise the error terms on the output of each polarization given in (4.6).

$$\begin{aligned} e_x &= 1 - |x_{out}|^2 \\ e_y &= 1 - |y_{out}|^2 \end{aligned} \quad (4.6)$$

While this optimisation criterion is known to function well in DP-QPSK, where both polarizations have constant power under noise free conditions, this approach is suboptimal for other constellations where the constant modulus condition is not met. For example, in the case of QAM16 modulation, some form of radially-directed algorithm is most commonly used to reduce the minimum CMA error term and therefore enhance convergence [100]. We therefore introduce a decision on the power in each input polarization to the equaliser such that a lower error term may be achieved. We also annotate this new algorithm as the polarization switched CMA or PS-CMA. This new error is described in (4.7):

$$\begin{aligned} (R_x \quad R_y) &= \begin{cases} (1 \quad 0), & \text{if } (|x_{out}| > |y_{out}|) \\ (0 \quad 1), & \text{if } (|x_{out}| \leq |y_{out}|) \end{cases} \\ e_x &= R_x - |x_{out}|^2 \\ e_y &= R_y - |y_{out}|^2 \end{aligned} \quad (4.7)$$

The taps of the four filters are then adapted as previously using the least mean squares algorithm given by (4.8):

$$\begin{aligned}
\mathbf{h}_{xx} &= \mathbf{h}_{xx} + \mu e_x \mathbf{x}_{in} \mathbf{x}_{out}^* \\
\mathbf{h}_{xy} &= \mathbf{h}_{xy} + \mu e_x \mathbf{y}_{in} \mathbf{x}_{out}^* \\
\mathbf{h}_{yx} &= \mathbf{h}_{yx} + \mu e_y \mathbf{x}_{in} \mathbf{y}_{out}^* \\
\mathbf{h}_{yy} &= \mathbf{h}_{yy} + \mu e_y \mathbf{y}_{in} \mathbf{y}_{out}^*
\end{aligned} \tag{4.8}$$

Where \mathbf{x}_{in} and \mathbf{y}_{in} are the input vectors to the equaliser on the x and y polarizations respectively, and the outputs of the equaliser x_{out} and y_{out} are given by (4.9):

$$\begin{aligned}
x_{out} &= \mathbf{h}_{xx}^H \mathbf{x}_{in} + \mathbf{h}_{xy}^H \mathbf{y}_{in} \\
y_{out} &= \mathbf{h}_{yx}^H \mathbf{x}_{in} + \mathbf{h}_{yy}^H \mathbf{y}_{in}
\end{aligned} \tag{4.9}$$

In order to ensure non-singular convergence the filter coefficients must be initialised in an optimal manner such that the central taps of the equaliser are aligned with the signal input polarization state. This is done by calculating the expected correlation between output powers for a series of test rotations of the initial tap weights, and select the angle of rotation for which the output powers are least correlated. We therefore initialise the four central taps to be of the form given by (4.2):

$$\begin{pmatrix} h_{xx} & h_{xy} \\ h_{yx} & h_{yy} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \tag{4.2}$$

Our instantaneous equaliser outputs will therefore be given by (4.3):

$$\begin{pmatrix} x_{out} \\ y_{out} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_{in} \\ y_{in} \end{pmatrix} \tag{4.3}$$

Our optimal filter coefficients are given by the state which has minimal correlation in output powers. This is equivalent to the condition in (4.4):

$$\begin{aligned}
\theta &= \arg \min_{\theta} \varepsilon^2 ; \quad \text{where:} \\
\varepsilon^2 &= \langle |x_{out}|^2 |y_{out}|^2 \rangle
\end{aligned} \tag{4.4}$$

where $\langle \cdot \rangle$ and $|\cdot|$ denote the expectation and modulus operators respectively. After some algebra, it may be shown that (4.4) is equivalent to (4.5):

$$\varepsilon^2 = a(1 - \cos(4\theta)) + b\sin(4\theta) + c(3 + \cos(4\theta)) \tag{4.5}$$

where:

$$a = \frac{1}{8} \langle |x_{in}|^4 + |y_{in}|^4 - 4|\text{Re}(x_{in}y_{in}^*)|^2 \rangle;$$

$$b = \frac{1}{2} \langle (|y_{in}|^2 - |x_{in}|^2)\text{Re}(x_{in}y_{in}^*) \rangle; \quad c = \frac{1}{4} \langle |x_{in}|^2 |y_{in}|^2 \rangle;$$

This one dimensional optimisation is straightforward to solve analytically for the optimal value of θ , or using a direct search from a set of test values of θ .

By finding the angle of polarization which satisfies the first derivative of correlation being zero and the second derivative being maximum, we may find the optimum angle analytically. A brief derivation is given in (4.6), with the analytical solution given in (4.7).

$$\frac{d\varepsilon^2}{d\theta} = 4((a - c)\sin(4\theta) + b\cos(4\theta)) = 0$$

$$\tan(4\theta) = \frac{b}{c - a}$$

$$\theta = \frac{1}{4} \tan^{-1} \left(\frac{b}{c - a} \right); \quad \text{or} \quad \theta = \pi - \frac{1}{4} \tan^{-1} \left(\frac{b}{c - a} \right) \quad (4.6)$$

$$\frac{d^2\varepsilon^2}{d\theta^2} = 16((a - c)\cos(4\theta) - b\sin(4\theta)) > 0$$

$$\theta = -\frac{\varphi}{4} = -\frac{1}{4} \sin^{-1} \left(\frac{b}{\sqrt{(a - c)^2 + b^2}} \right) \quad (4.7)$$

While this method gives an exact solution to (4.4), it may be preferable to find an approximate solution with a parameter sweep to avoid the complex arithmetic which is required to calculate (4.7).

4.5 Comparison of the PS-CMA with Generic Equalisers

Another commonly used equaliser algorithm is the decision directed algorithm, which has the advantage of being a generic equaliser which may be applied to any modulation format [46]. While a decision directed LMS (DD-LMS) algorithm may be applied to any modulation format in any number of dimensions, some pre-conditioning of the filter coefficients with another algorithm is normally required

such that convergence may be reliably attained [10]. This may be demonstrated with a comparison of the abilities of various algorithms to recover an arbitrary polarization rotation. We therefore compared the performance of the PS-CMA with and without the initialisation algorithm to the DD-LMS algorithm.

The decision-directed algorithm uses an error term given by (4.16):

$$\begin{aligned} e_x &= D_x(x_{out}) - x_{out} \\ e_y &= D_y(y_{out}) - y_{out} \end{aligned} \quad (4.16)$$

$$D(u) = \begin{pmatrix} D_x(u) \\ D_y(u) \end{pmatrix} = \underset{s}{\operatorname{argmin}} |u - s|, \text{ where } s \in S_{PS-QPSK}$$

where D is a four dimensional minimum Euclidean distance hard decision, as defined above.

The input signal was noise loaded to 5.8 dB E_b/N_0 . This may be found to be commensurate with a BER of 10^{-3} , according to the combination of equations (4.2) and (4.3) when ‘*conventional*’ mapping is applied. The signal was then rotated to a variety of polarization using a Jones matrix such that $-\pi/2 \leq \phi < \pi/2$ and $-\pi \leq \theta < \pi$, thus ensuring that the Poincare sphere is evenly covered. The Jones matrix used is described in (4.17).

$$\mathbf{J} = \begin{pmatrix} \cos(\theta) & \sin(\theta)e^{j\phi} \\ -\sin(\theta)e^{-j\phi} & \cos(\theta) \end{pmatrix} \quad (4.17)$$

The signal was then equalised before bit error counting and BER calculation. Q-factor was then calculated according to [120] from the observed BER, and this was compared to the ideal Q-factor corresponding to a BER of 10^{-3} to give the Q Penalty. For both the PS-CMA and DD-LMS with PS-CMA conditioning, equaliser performance was found to be ambivalent to the input state of polarization, with Q-factor penalty being uniformly low. Figure 4.7 shows the sensitivity of the considered equalisers to the input state of polarization for the DD-LMS equaliser with no conditioning.

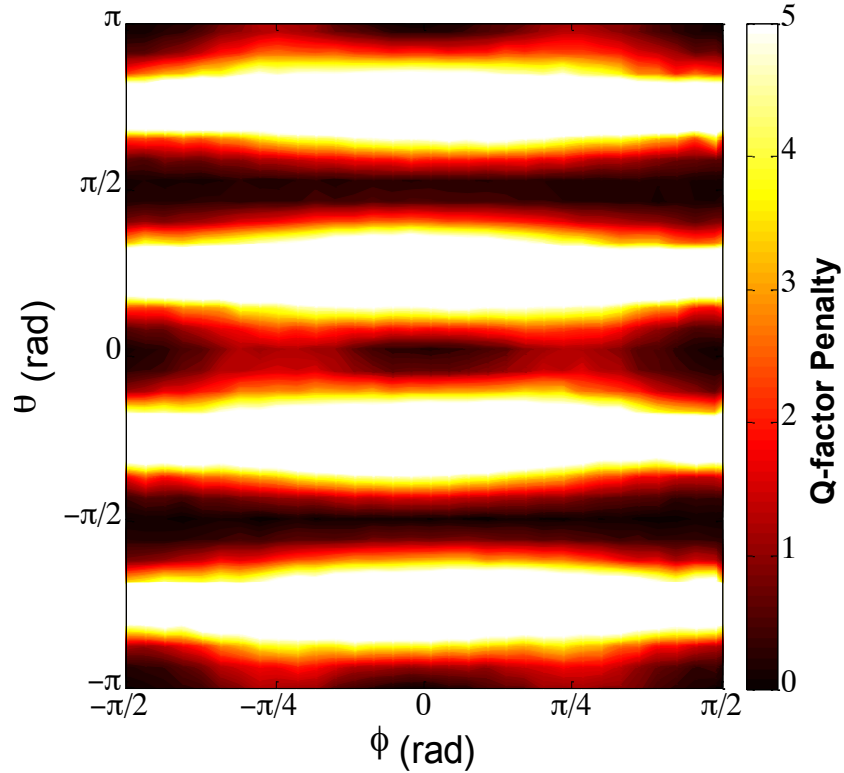


Figure 4.7 - Input polarization sensitivity of the DD-LMS with PS-QPSK modulation. Q-factor penalty in dB is plotted against the two angular parameters in the Jones matrix, with dark colour representing a low penalty and light colour representing a high penalty.

We note from Figure 4.7 that the DD-LMS equaliser is highly sensitive to the input polarization state, and converges well for relatively few states. The PS-CMA is therefore a useful algorithm, both in its own right and for conditioning filter coefficients such that the DD-LMS may be used.

The input polarization sensitivity of the PS-CMA was characterised with and without the initialisation procedure described above. The mean Q-factor penalty was calculated over the dimension ϕ in each case and the results presented in Figure 4.8. These results demonstrate the effectiveness of the initialisation procedure in eliminating the problem of singular mal-convergence, which results in a large Q-factor penalty. A high penalty is seen for values of θ of $\pm\pi/4$ and $\pm3\pi/4$, which correspond to angles whereby the energy of each signal polarization is evenly distributed onto the two receiver polarizations. This behaviour is similar to that noted for the DD-LMS equaliser shown in Figure 4.7.

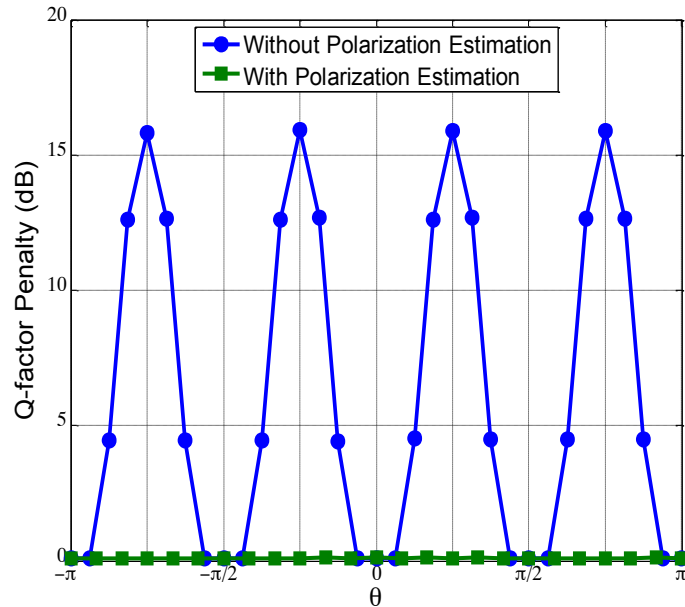


Figure 4.8 - Polarization sensitivity of the PS-CMA with and without initialisation algorithm. Mean Q-factor penalty is plotted against polarization angle.

The DP-CMA with PS-QPSK modulation was found to provide only a degenerate solution with 1 of 3 transmitted bits being discarded. As the DP-CMA equaliser attempt to form a constant modulus with both polarization input signals, the lowest error will be reached when both input polarizations are summed and the output polarizations are equal. Here we noted that regardless of the input state of polarization, in all cases all information contained in the bit which encodes polarization is lost.

The inherent mal-convergence problem with the DP-CMA is illustrated by Figure 4.9. When the PS-CMA equaliser is used, the constellation is fully recovered, and we clearly see the four constellation points belonging to the QPSK constellation on the polarization plotted, and the point at the origin which occurs when the QPSK signal is being transmitted on the other polarization. When the DP-CMA is used, both polarization outputs give the same constellation, which contains only the points which belong to the QPSK subset, and all information encoded on polarization is lost.

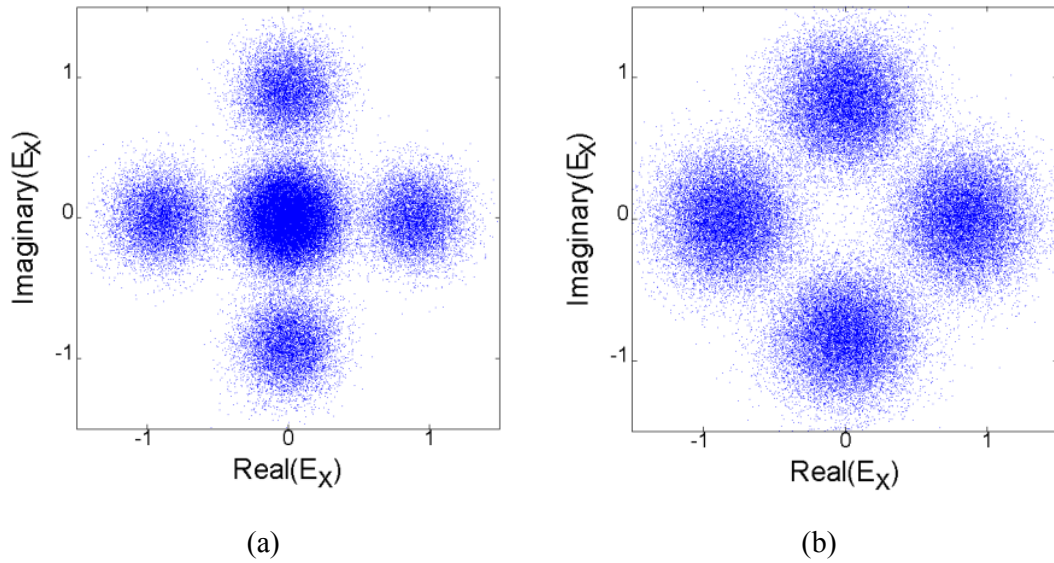


Figure 4.9 - Constellation diagrams showing a single output polarization of PS-QPSK modulation when equalised with the PS-CMA equaliser (a); and PS-QPSK modulation when equalised with the DP-CMA equaliser (b).

Noise loaded to 5.8 dB E_b/N_0 , for illustrative purposes.

4.6 Characterisation of PS-CMA Performance

An essential capability of any practical equaliser is the ability to operate effectively in the presence of polarization dependant loss (PDL). To characterise the performance of the equaliser in the presence of PDL we took the transmitted optical signal, applied loss at a particular polarization orientation and then noise loaded to 5.8 dB E_b/N_0 . We considered the orientation of the loss as a polarization which is circularly rotated with respect to the signal, that is, θ was varied between $-\pi$ and π , while ϕ remained zero. The signal is then rotated again with respect to the polarization axes of the receiver such that the input state of polarization to the receiver is randomised. The mean Q-factor penalty with respect to the SNR limited optimum was then calculated over the entire polarization space. The results of this analysis are presented below in Figure 4.10.

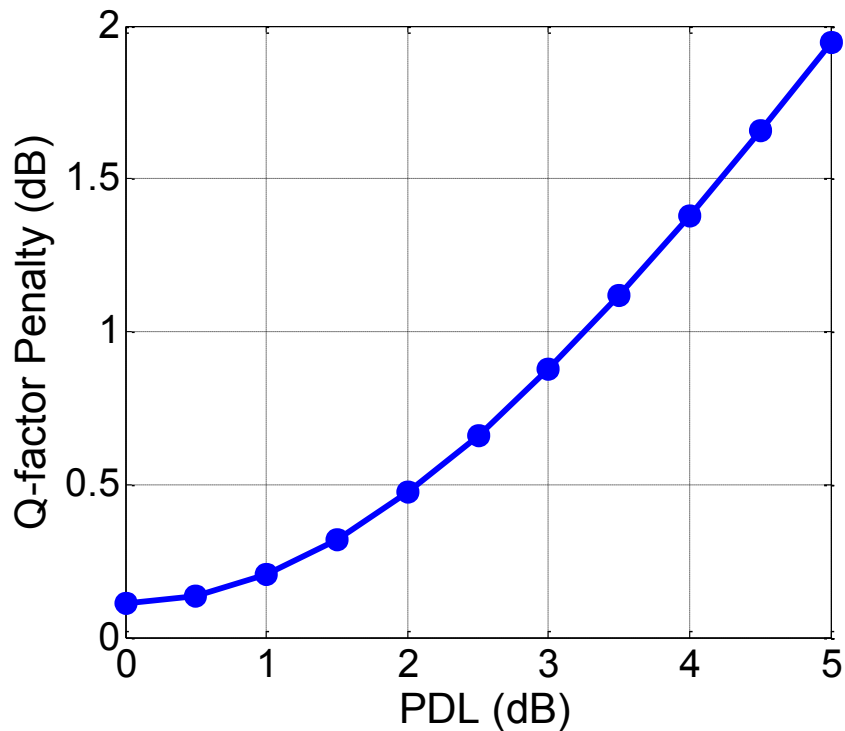


Figure 4.10 - Performance of the PS-CMA with PS-QPSK modulation in the presence of PDL. Q-factor penalty in dB is plotted against the applied PDL in dB.

From Figure 4.10 we note that 3 dB of PDL results in a Q-factor penalty of approximately 1 dB. It is noted that while the Q penalty of the received signal varies with the amount of PDL, this does not increase the likelihood of the singularity problem which we described earlier for levels of PDL up to 5 dB, unlike when DP-QPSK modulation is used with the DP-CMA equaliser. While some improvements in performance are possible when using the DP-CMA with DP-QPSK modulation by constraining the filters [121], [122] mal-convergence probability due to PDL is always non-zero, whilst the PS-CMA with PS-QPSK always converges correctly.

Practical transmission links will always suffer from polarization mode dispersion (PMD), which in the past was a performance limiting distortion. It has been previously demonstrated that the DP-CMA may compensate for arbitrarily large amounts of PMD with DP-QPSK modulation given a sufficiently large number of taps. We analysed the tolerance to differential group delay (DGD) of a 25 tap PS-CMA with PS-QPSK modulation, with the number of taps chosen to enable up to 6 symbol periods of DGD to be equalised. Simulations were performed by circularly rotating the polarization of the transmitted optical signal, applying a delay in the

frequency domain to attain the desired level of DGD, and then rotating the signal polarization back to its original state. The signal was then noise loaded as before to 5.8 dB E_b/N_0 , before equalisation and error counting.

The equaliser was found to be ambivalent to the orientation of the applied DGD, being able to recover the input signal without penalty for DGD of up to 6 symbols. This is in contrast to DP-QPSK with the DP-CMA which is known to suffer mal-convergence when the energy is split equally between the two principal states of polarization.

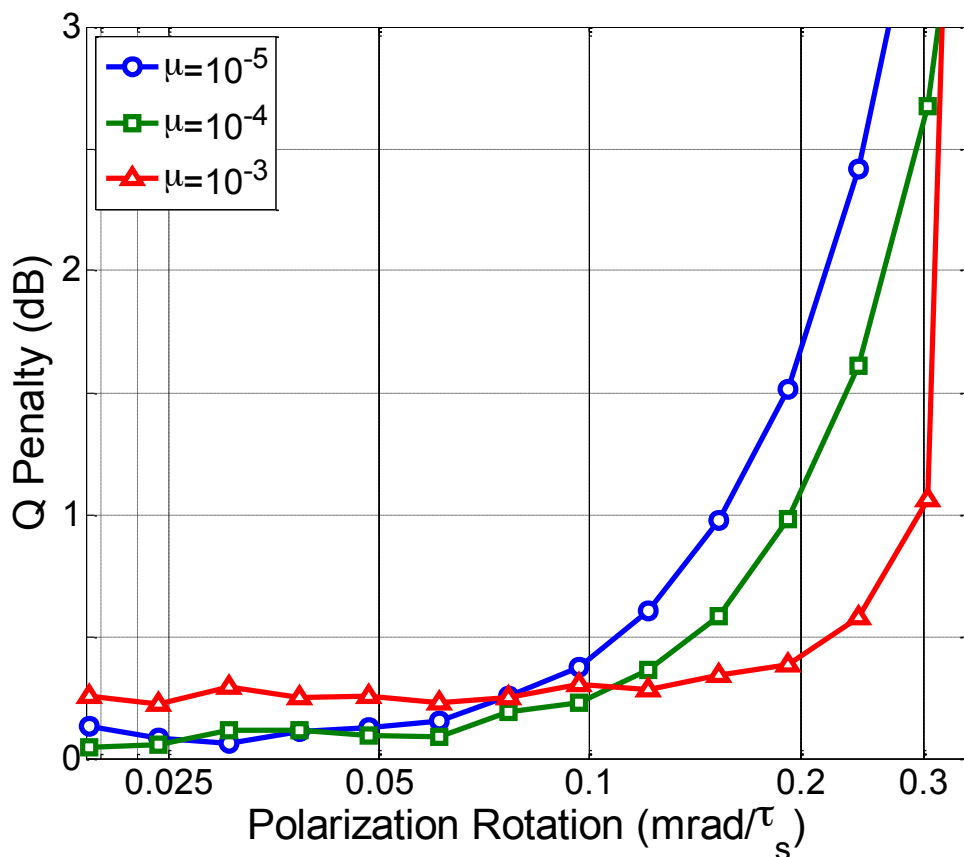


Figure 4.11 - Performance of the PS-CMA with conventionally coded PS-QPSK modulation in the presence of time varying polarization rotation.

Due to the time-varying birefringence of optical fibre, another important characteristic of a digital equaliser for coherent optical communication is the ability to track the time varying state of polarization at the input of the receiver. To measure the performance of the receiver in this respect, we rotated the transmitted signal by a Jones matrix with a time varying circular rotation, such that ϕ remains zero and θ is increased at a constant rate to produce a rotation with constant angular frequency.

The signal was then noise loaded to 5.8 dB E_b/N_0 and equalised with the PS-CMA prior to error counting. Performance was measured by Q-factor against both the angular frequency of polarization rotation and the PS-CMA convergence parameter μ . The results of this simulation are presented in Figure 4.11. The same simulation was also performed utilising differential coding as described previously, with the results presented in Figure 4.12.

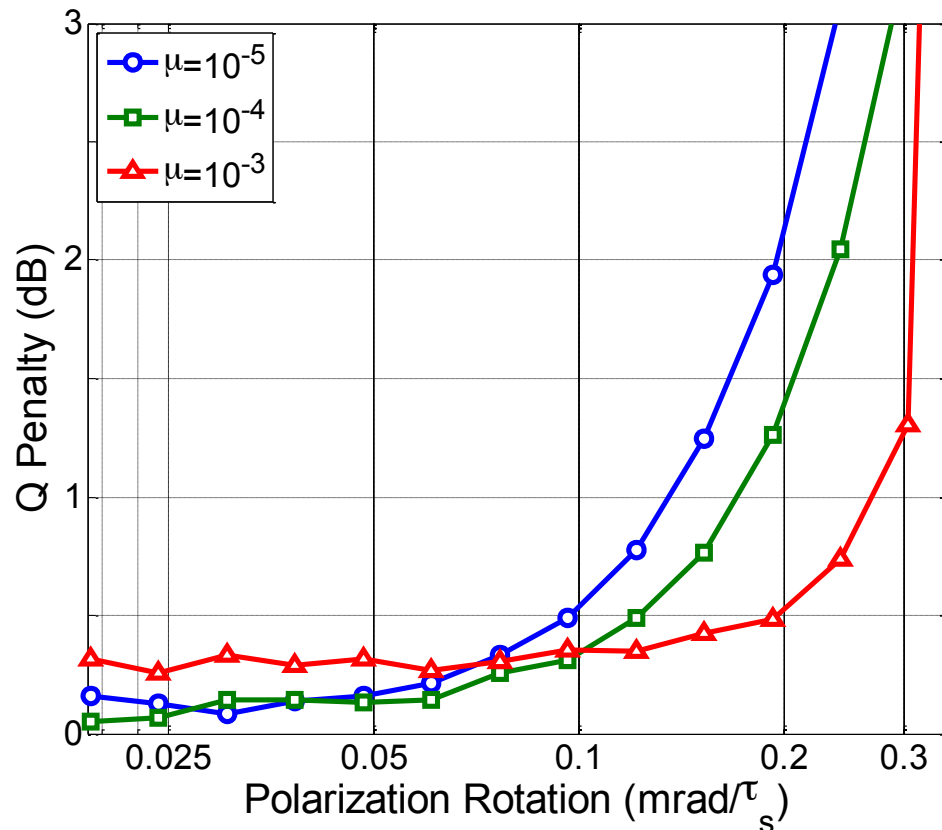


Figure 4.12 - Performance of the PS-CMA with differentially coded PS-QPSK modulation in the presence of time varying polarization rotation.

It is noted from Figure 4.11 and Figure 4.12 that an increased convergence parameter μ enables a faster polarization rotation to be tracked, at the expense of a reduction in receiver sensitivity. It is also noted that for both coding schemes, a polarization rotation frequency of approximately 0.1 mrad per symbol period may be tracked for a penalty in performance of approximately 0.5 dBQ. Despite the fact that the differential coding scheme presented in Figure 4.12 has differential coding on polarization as well as phase, we note that there is no appreciable improvement in polarization tracking ability in comparison with conventional coding. This indicates that the polarization tracking capability of this system is limited by the convergence

speed of the equaliser rather than burst errors caused by polarization slips. This also indicates that in a real system, polarization tracking will be significantly worse than is described here, as the equaliser will be updated at the ASIC clock rate (normally in the region of 500 MHz) rather than the symbol rate as is assumed here.

4.7 Comparison with Other Published PS-QPSK Algorithms

To-date, only two papers have directly addressed the issue of fully blind equalisation of PS-QPSK. The algorithm used by Nelson et al in [115] is a normalised version of the PS-CMA derived above, although the initialisation algorithm was not used, and so the singularity issue will still be present. This work was done in parallel to and was published shortly after our research, presented in [13] and [15]. An alternative equaliser algorithm was presented in [14] while our work was under review. While the algorithm presented in [14] uses a different control surface to the PS-CMA described above, the operation is remarkably similar. The only significant difference between the two algorithms is the initialisation procedure described in section 4.4, which mitigates the possibility of singular mal-convergence. To compare the performance of the two algorithms, we have simulated the back-to-back performance of both algorithms after a polarization rotation. The simulation was identical to that previously described for Figure 4.8.

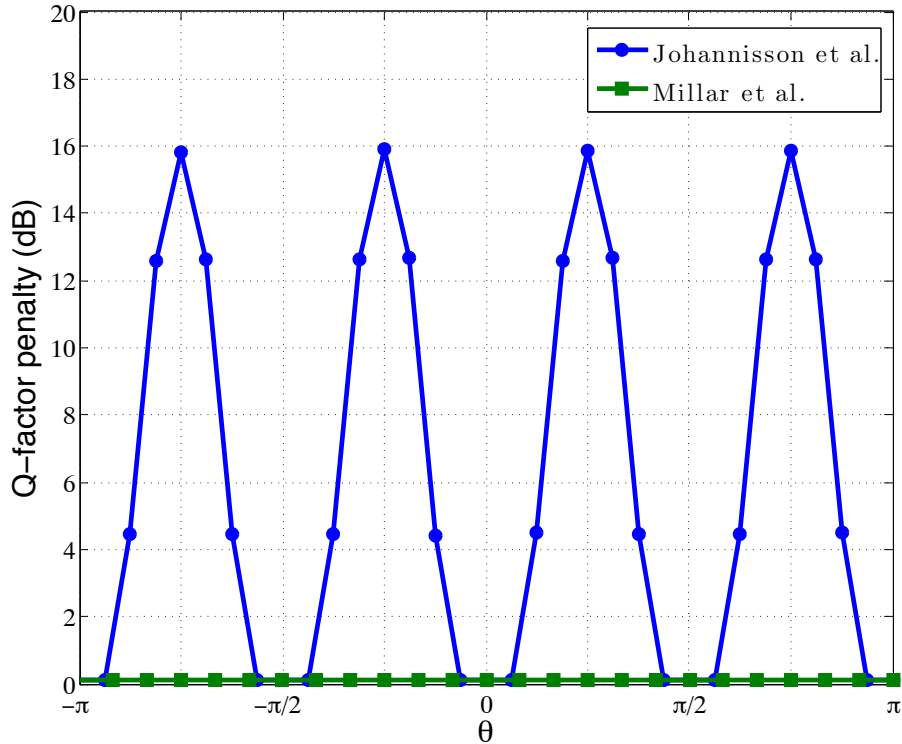


Figure 4.13 - Q-factor penalty as a function of input polarization state for the equaliser algorithm as described in [14] (Johannisson et al.) compared with that described in section 4.4 (Millar et al.).

We note from Figure 4.13 that while the minimum Q-factor penalty of the equalisers is broadly similar in the optimum case, the algorithm described in [14] (Johannisson et al.) suffers from singular mal-convergence resulting in a high Q-factor penalty. A high penalty is seen for values of θ of $\pm\pi/2$ and $\pm3\pi/2$, which correspond to angles whereby the energy of each signal polarization is evenly distributed onto the two receiver polarizations. This result is in agreement with the previous analysis of the performance of the PS-CMA with and without the initialisation algorithm presented in Figure 4.8 and demonstrates the improvement in robustness which may be achieved with the initialisation algorithm.

4.8 Application of PS-QPSK to 100 GbE WDM Systems

High-speed optical transmission systems such as the proposed PS-QPSK system are almost universally utilised in combination with wavelength division multiplexing (WDM). The current standards for first generation deployment of 100 GbE systems are for 112 Gb/s per wavelength DP-QPSK modulation over a 50 GHz WDM grid, yielding a spectral efficiency of 2 b/s/Hz after overheads [11]. An interesting comparison is therefore to examine the receiver sensitivity of both PS-QPSK and DP-QPSK WDM systems at 112 Gb/s with a spectral efficiency of 2 b/s/Hz, corresponding to symbol rates of 37.3 GBd and 28 GBd respectively. For completeness we compare conventional (which corresponds to non-differential Gray coding for DP-QPSK) and differential coding for both modulation formats. Although the use of PS-QPSK will require use of a broader spectrum than that used for DP-QPSK, we may use electrical filtering to reduce the effects of linear cross-talk introduced by WDM. In Figure 4.14 we have compared the penalty in required OSNR to achieve a BER of 10^{-3} for differing analogue electrical bandwidths at the transmitter and receiver. The transmitted signal was a modulated impulse train, filtered using a 5th order low-pass Bessel filter, and then combined with two additional channels spaced at 50GHz using an ideal colourless power combiner. The signal was then noise loaded and detected with an ideal coherent receiver. The signal was then filtered again with a second identical 5th order low-pass Bessel filter, resampled to 2 samples per symbol and then equalised with the PS-CMA for PS-QPSK or the DP-CMA for DP-QPSK to remove the response of the analogue electrical filtering. After symbol estimation and BER calculation, the required OSNR was calculated. By varying the bandwidth of the two electrical filters, we determine the minimum required OSNR. The results of this simulation are presented in Figure 4.14.

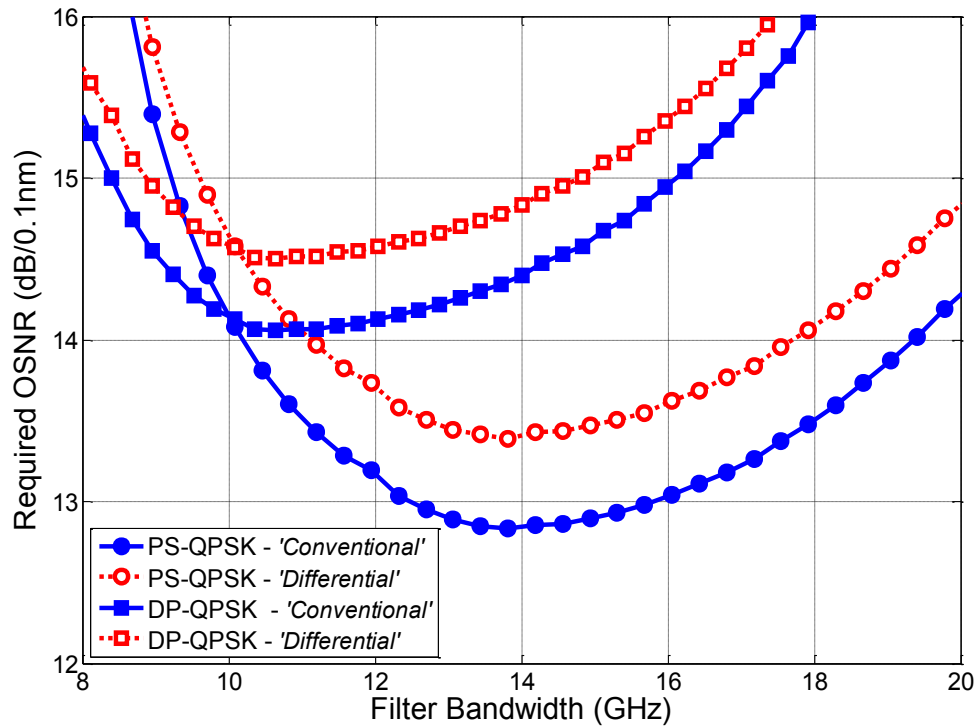


Figure 4.14 - Back-to-back comparison of required OSNR to achieve a BER of 10^{-3} for PS-QPSK and DOP-QPSK. Both systems are three channels at 112 Gb/s (7% FEC overhead) over 50 GHz WDM grid.

From Figure 4.14, we note that the required OSNR for PS-QPSK is slightly more than 1 dB lower than that required for DP-QPSK. This relative penalty is maintained when differential coding is introduced. For lower electrical bandwidths in all cases, a penalty results from the signal filtering, while for higher electrical bandwidths, a penalty results from linear cross-talk between adjacent WDM channels.

The simulation was then repeated for a 3 channel WDM system of 124.8 Gb/s over a 50 GHz grid. The increase in FEC overhead results in an increase in baud rate to 41.6 GBd for PS-QPSK and 31.2 GBd for DP-QPSK. In this case the FEC overhead was 20%, with a BER limit of 2×10^{-2} . The results are presented in Figure 4.15.

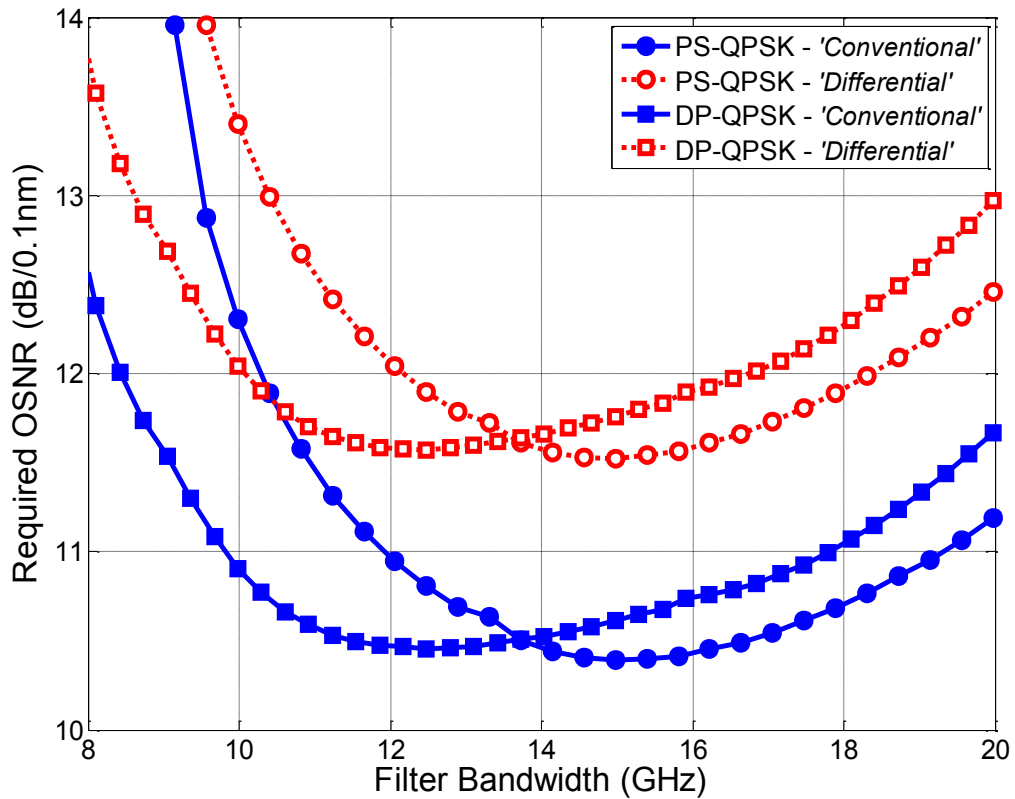


Figure 4.15 - Back-to-back comparison of required OSNR to achieve a BER of 2×10^{-2} for PS-QPSK and DP-QPSK. Both systems are three channels at 124.8 Gb/s (20% FEC overhead) over a 50 GHz WDM grid.

From Figure 4.15, we notice that the improvement in required OSNR is reduced to less than 0.1dB when a higher FEC overhead of 20% is used, due to the broader spectrum of the signals. Again, we note that in the low electrical bandwidth regime there is a significant penalty from filtering, while in the high bandwidth regime there is significant penalty from cross-talk. However, the benefit from using PS-QPSK is reduced from 0.5 dB (at a BER of 2×10^{-2}) to less than 0.1 dB by cross-talk. While this negates much of the benefit of PS-QPSK, this may be outweighed in some circumstances by the practical benefit due to the increased robustness of the PS-CMA equaliser.

4.9 Summary

In this chapter, we have examined the efficiency of several bit mapping schemes for PS-QPSK and used them to characterise the performance of novel DSP algorithms for this modulation format.

A novel polarization-switched CMA equaliser was proposed and was found to have comparable performance to the DP-CMA equaliser, but does not suffer from degenerate mal-convergence with certain input polarizations states or PDL up to 5 dB. The PS-CMA may therefore be said to be more robust than the DP-CMA without the use of training sequences or significant additional algorithmic complexity.

A comparison was performed between PS-QPSK and DP-QPSK for 100 GbE WDM systems over a 50 GHz frequency grid. PS-QPSK was found to enable a gain in receiver sensitivity of 1 dB in required OSNR when compared with standard DP-QPSK and a 7% FEC overhead.

In the next chapter, we will examine the nonlinear transmission performance of PS-QPSK. Generation of signals is discussed, and long-haul transmission of 40 Gb/s WDM PS-QPSK is compared to DP-QPSK at equal spectral efficiencies. We will determine if the theoretical benefit of PS-QPSK which is apparent in the simulated systems presented in this chapter can be maintained in experimental systems.

5 Generation and Long-Haul Transmission of Polarization-Switched QPSK at 42.9 Gb/s

5.1 Abstract

In this chapter, we demonstrate for the first time generation and transmission of polarization-switched QPSK (PS-QPSK) signals at 42.9 Gb/s. Long-haul transmission of PS-QPSK is experimentally investigated in a recirculating loop and compared with transmission of dual-polarization QPSK (DP-QPSK) at 42.9 Gb/s per channel. A reduction in the required OSNR of 0.7 dB was found at a BER of 3.8×10^{-3} , resulting in an increase in maximum reach of more than 30% for a WDM system operating on a 50 GHz frequency grid. The maximum reach of 13640 km for WDM PS-QPSK is, to the best of our knowledge, the longest distance reported for 40 Gb/s WDM transmission, over an uncompensated link, with standard fibre and amplification.

5.2 Introduction

Coherent detection, combined with digital signal processing (DSP), has led to recent increases in capacity [109], reach [123], and spectral efficiency [124]. While much research has been devoted to generating, processing and transmitting spectrally-efficient modulation formats such as high-level quadrature amplitude modulation (QAM) [100], these modulation formats rely on a regular 4-dimensional lattice constellation design. However, it has been recently shown that these modulation formats are not optimal for the optical channel in terms of the *asymptotic power efficiency* [111], and some work has been undertaken to determine the optimal modulation format for a 4-dimensional additive White Gaussian noise channel [34], [111].

This research has led to a variety of new modulation formats being proposed, with various degrees of complexity and difficulty of realisation. A format which has attracted interest is polarization-switched quadrature phase shift keying (PS-QPSK) [14], [34], [111], [116], [117], [125]. This format transmits a symbol, on one of two orthogonal polarizations, with one of four equally spaced phase levels from a QPSK constellation, such that the resulting symbol carries 3 bits of information.

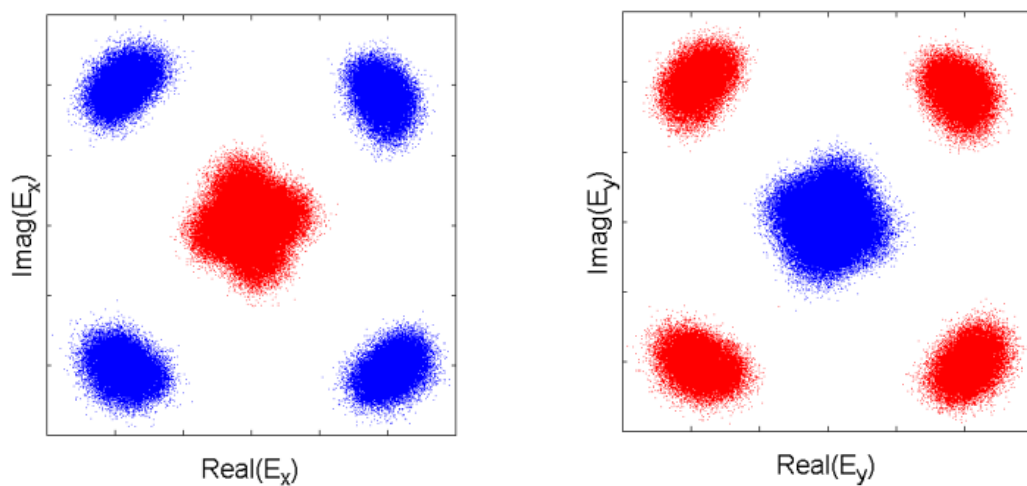


Figure 5.1 - Constellation diagrams showing two orthogonal linear polarizations of an experimentally generated PS-QPSK signal.

This is illustrated with a pair of experimental constellations in Figure 5.1, where blue points denote a QPSK symbol which has been transmitted on the x-polarization, while red dots denote a QPSK symbol transmitted on the y-polarization. Whilst this

modulation format has a lower available spectral efficiency than DP-QPSK, the gain in noise tolerance is as much as 1.76 dB [111] at equal bit rates and asymptotically high optical signal to noise ratio (OSNR). For the bit-error-rate (BER) values combined with modern forward error correction (FEC) codes, an improvement of 1 dB at a BER of 10^{-3} and of 0.55 dB at a BER of 10^{-2} is theoretically achievable. Due to the significant benefit in noise tolerance over DP-QPSK which has become the standard modulation format for 100 GbE technology [11], there has been some interest recently in the transmission properties of PS-QPSK [116], [117], [125]. For this research we used an easily realisable technique to experimentally generate PS-QPSK, without the use of either a four-dimensional modulator or custom-made photonic integrated circuits. PS-QPSK was then characterised for long-haul WDM transmission and compared to DP-QPSK.

5.3 PS-QPSK Generation and Experimental Setup

To evaluate the achievable benefits of employing the PS-QPSK modulation format over DP-QPSK, we first measured the OSNR tolerance and maximum reach of both modulation formats. The experiments were conducted at the constant bit rate of 42.9 Gb/s, corresponding to 14.3 Gbaud for PS-QPSK and 10.725 Gbaud for DP-QPSK. In WDM transmission, both formats were transmitted over a 50 GHz frequency grid, with spectral efficiency of 0.8 b/s/Hz in both cases.

The PS-QPSK format was generated as follows (Figure 5.2(b)). First a triple Mach-Zehnder modulator (MZM) was used to modulate CW light from an external cavity laser (ECL) to generate a single polarization QPSK sequence. The applied data pattern was a pseudo-random bit sequence (PRBS) of length $2^{15}-1$, where the PRBS was decorrelated by half of the pattern length between the in-phase and quadrature signal components. Polarization switching was then applied to the signal by passively 50:50 splitting the QPSK signal and intensity modulating each arm, symbol-synchronously, with two MZMs. The two intensity modulators were driven by DATA and $\overline{\text{DATA}}$ respectively from the pattern generator. The effect of this configuration being that only one intensity modulator was transmitting during each symbol period. The two arms were then tuned to be orthogonally polarized using polarization controllers, before entering a polarization beam combiner. Any residual

symbol timing difference, due to different path lengths in the polarization switching stage, was compensated with a variable optical delay line in one arm.

DP-QPSK was generated using a similar method except that, after the 50:50 splitter, the QPSK signal was polarization-multiplexed by decorrelating the signal polarizations in each arm of a passive delay-line stage (Figure 5.2(a)). The effective delay between the QPSK signals in each polarization was 24 symbols.

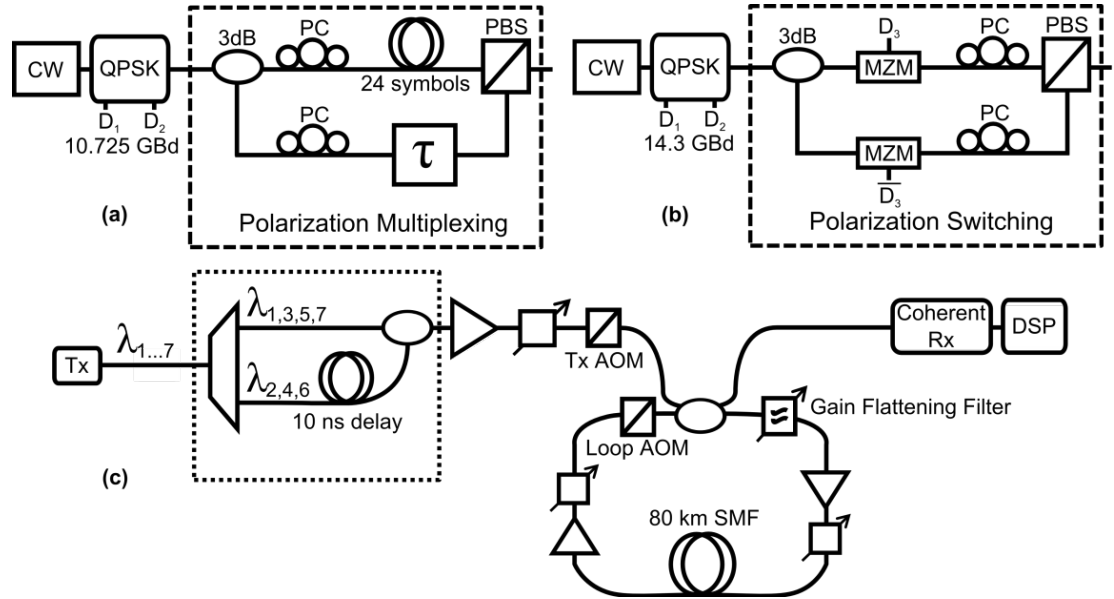


Figure 5.2 - Experimental set-up to generate and transmit 42.9 Gb/s PS-QPSK (14.3 Gbaud) and DP-QPSK (10.725 Gbaud).

In order to generate a 7-channel WDM comb, two alterations to the above configuration were introduced. Firstly, a 50 GHz WDM comb was created by combining CW light from six temperature- and current-controlled DFB lasers. An ECL with a linewidth of 100 kHz was used for the central channel. The comb was bulk modulated, and an interleaver with a channel spacing of 50 GHz was used to separate alternate channels, which were then decorrelated by 10 ns before being recombined with a 3-dB fibre coupler, as described in (A H Gnauck et al. 2011).

To investigate the transmission performance, a single-span recirculating fibre loop was used. An EDFA followed by a variable optical attenuator (VOA) was used to set the launch power to an acousto-optic modulator controlled recirculating fibre loop (Figure 5.2(c)), as described in [126]. The loop span consisted of 80.24 km of standard single-mode fibre (SMF), with a loss of 15.4 dB and total chromatic

dispersion of 1347 ps/nm at 1554 nm. Wavelength-dependant gain of the loop EDFAs was equalised using gain flattening filters.

After the desired number of recirculations, the signal was detected with a phase and polarization diverse coherent receiver. The frequency offset between the signal and LO lasers typically measured as being less than 1 GHz. Digitisation was then performed using a real-time digital sampling oscilloscope (DSO) with 8 physical bits of resolution and 50 GSa/s. The captured waveforms were subsequently processed offline using Matlab. When processing DP-QPSK, we used the linear processing methods, described in detail in [12].

5.4 Digital Signal Processing for Experimentally Generated PS-QPSK Signals

The captured digital signal was first de-skewed, normalised to unit power per polarization and re-sampled to 2 Samples/symbol. Equalisation was then performed using a polarization-switched constant modulus algorithm (PS-CMA) equaliser with least-mean squares (LMS) updating, described in detail in the previous chapter, and originally in [13]. This equaliser utilises a decision on the relative power in each output polarization from the equaliser. The error term may be described by the following pseudo-code (5.1):

$$\begin{aligned} (R_x \quad R_y) &= \begin{cases} (1 \quad 0), & \text{if } (|x_{out}| > |y_{out}|) \\ (0 \quad 1), & \text{if } (|x_{out}| \leq |y_{out}|) \end{cases} \\ e_x &= R_x - |x_{out}|^2 \\ e_y &= R_y - |y_{out}|^2 \end{aligned} \tag{5.1}$$

The taps of the four filters are then adapted as previously described using the least mean squares algorithm given by (5.2):

$$\begin{aligned} \mathbf{h}_{xx} &= \mathbf{h}_{xx} + \mu e_x \mathbf{x}_{in} \mathbf{x}_{out}^* \\ \mathbf{h}_{xy} &= \mathbf{h}_{xy} + \mu e_x \mathbf{y}_{in} \mathbf{x}_{out}^* \\ \mathbf{h}_{yx} &= \mathbf{h}_{yx} + \mu e_y \mathbf{x}_{in} \mathbf{y}_{out}^* \\ \mathbf{h}_{yy} &= \mathbf{h}_{yy} + \mu e_y \mathbf{y}_{in} \mathbf{y}_{out}^* \end{aligned} \tag{5.2}$$

Where \mathbf{x}_{in} and \mathbf{y}_{in} are the input vectors to the equaliser on the x and y polarizations

respectively, and the outputs of the equaliser x_{out} and y_{out} are given by (5.3):

$$\begin{aligned}x_{out} &= \mathbf{h}_{xx}^H \mathbf{x}_{in} + \mathbf{h}_{xy}^H \mathbf{y}_{in} \\y_{out} &= \mathbf{h}_{yx}^H \mathbf{x}_{in} + \mathbf{h}_{yy}^H \mathbf{y}_{in}\end{aligned}\tag{5.1}$$

This equaliser has an attractive practical advantage over that which is used for dual-polarization QPSK. The PS-CMA equaliser when initialised with the algorithm described in [13] (and in detail in the previous chapter) does not suffer from degenerate mal-convergence (whereby both output polarizations of the equaliser converge to the same input polarization), regardless of the input polarization. This is due to the fact that the two switched QPSK tributaries are orthogonal and the equaliser may be initialised to avoid singularity.

To estimate the intradyne frequency offset, the PS-QPSK symbol sequence was reduced to a QPSK symbol sequence at the output of the receiver. This was done by making relative decisions on the energy in each symbol, the polarization with higher energy in each symbol was determined to contain the QPSK phase information, and was extracted. The resultant QPSK sequence was raised to the 4th power to remove the modulation. The offset was then determined by finding the peak power in the FFT of the signal [53].

Carrier phase estimation was performed after removing the polarization modulation. This was done by selecting one sample per symbol based on which polarization has more energy. Carrier phase was then recovered for our reduced QPSK sequence using the Viterbi & Viterbi algorithm [55].

The transmitted symbol sequence was determined by correlating the received symbol sequence with the PRBS data, and calculating the delays and phase rotations in the transmitter and channel. The transmitted and received symbol sequences were then transformed into three bit sequences each using the bit-mapping described in [111], which were then compared to calculate the BER.

5.5 Back-to-Back and WDM Transmission Results at 42.9 Gb/s

First, the receiver noise sensitivity was measured for DP- and PS-QPSK using additional noise loading at the receiver. The results of this back-to-back measurement are shown in Figure 5.3, together with the theoretical SNR limit as derived in [34].

From Figure 5.3(a) we note an implementation penalty of 0.8 dB for single channel PS-QPSK, with an excess WDM implementation penalty of 0.2 dB at a BER of 3.8×10^{-3} . This is compared to the DP-QPSK measurements shown in Figure 5.3(b), where we see an implementation penalty of 0.9 dB for single channel with negligible excess WDM implementation penalty. We note that in absolute terms, the required OSNR of WDM PS-QPSK is 8.1 dB, compared to 8.8 dB for DP-QPSK.

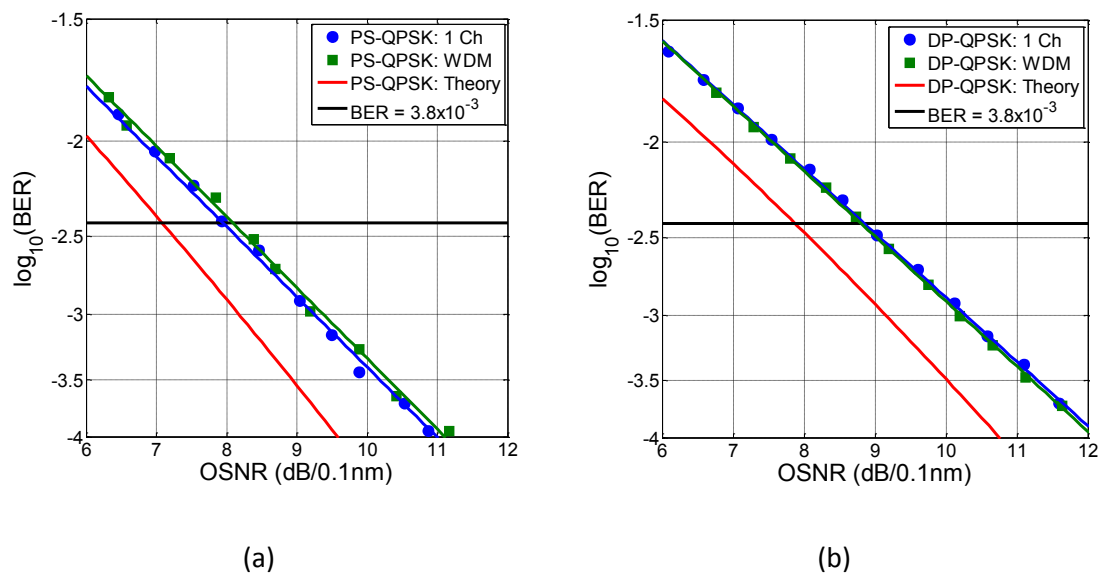


Figure 5.3 - Back-to-back measurements. Single-channel and WDM receiver OSNR sensitivity for (a) PS-QPSK and (b) DP-QPSK.

The transmission performance was then experimentally measured for a 7-channel WDM system using the recirculating loop. The launch power per channel was varied between -13 and 3 dBm, to determine the variation in maximum reach with launch power at the BER limit of 3.8×10^{-3} . The resulting comparison between PS-QPSK and DP-QPSK at 42.9 Gb/s is presented in Figure 5.4, with a polynomial fit for each curve as described in [75].

It can be seen from Figure 5.4 that in the low power, linear transmission regime (less than -7 dBm per channel), for a given reach a reduction in launch power of approximately 1dB per channel is possible for PS-QPSK, agreeing with the back-to-back receiver sensitivity results in Figure 5.3. For both modulation formats, the optimum launch power was found to be approximately -3.5 dBm per channel. The maximum reach of PS-QPSK was found to be 170 recirculations corresponding to 13,640 km; this may be compared to a maximum reach of 129 recirculations for DP-QPSK, corresponding to 10,350 km.

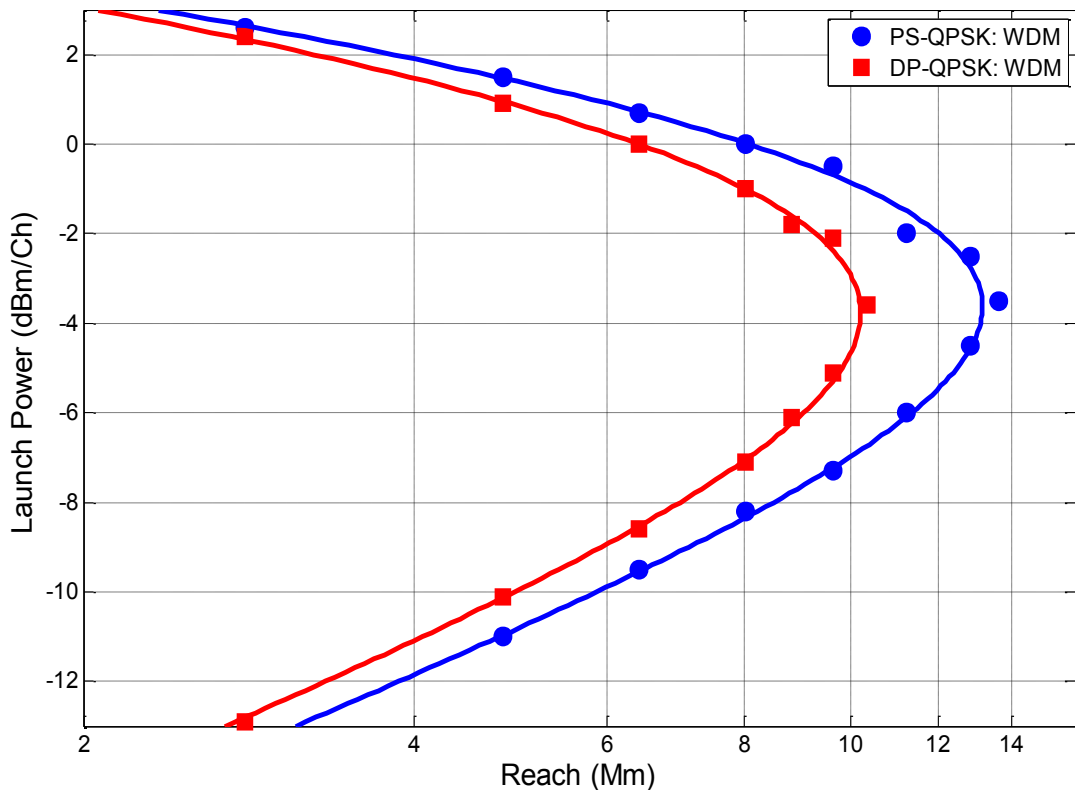


Figure 5.4 - Transmission performance of 42.9 Gb/s PS-QPSK and DP-QPSK compared for a 7 channel WDM system on a 50 GHz frequency grid. Maximum reach is compared for a BER of 3.8×10^{-3} .

The use of PS-QPSK rather than DP-QPSK modulation therefore enabled an increase in maximum reach of more than 30%. In the high power, highly nonlinear transmission region (launch power greater than 0 dBm per channel), we note that the improvement in performance available from PS-QPSK is reduced in comparison to the linear regime. This reduction in improvement was due to the high levels of nonlinear phase noise present.

5.6 Summary

In this chapter, we have presented the first experimental measurements of the transmission performance of PS-QPSK at 42.9 Gb/s using a simple and easily realisable generation technique. The implementation penalty for PS-QPSK modulation was found to be less than 1 dB at a BER of 3.8×10^{-3} . This transmitter was then used to perform a characterisation of transmission performance for PS-QPSK, comparing PS-QPSK and DP-QPSK at 42.9 Gb/s over a 50 GHz frequency grid. An improvement in launch power margin (the range of launch powers for which the FEC limit may be maintained) of greater than 1 dB was found in all cases. The optimum launch power for both modulation formats was found to be -3.5 dBm per channel, while maximum reach was increased by more than 30% from 10350 km to 13640 km. At the time of publication [15], these results represented the longest distance 40 Gb/s WDM transmission achieved over an uncompensated link, with standard fibre and amplification.

6 Conclusions and Topics for Further Research

6.1 Abstract

In this chapter, we describe proposed future research topics in the field of DSP algorithms for coherent detection. First, we discuss performance bounds on the capability of digital backpropagation for coherent systems. We then discuss DSP algorithms for modulation formats which are optimal in four dimensions. Extending the filter initialisation algorithm presented in chapter 4 to conventional modulation formats such as DP-QPSK and DP-QAM16 is then proposed. Finally we draw conclusions on each part of the research, and on the larger conclusions to be drawn from this research project as a whole.

6.1.1 Boundaries on Performance Benefits resulting from Nonlinearity Compensation

While we have performed a detailed investigation into the potential performance benefits of digital backpropagation in chapter 3, the backpropagation algorithm is not known to be optimal in the presence of random distortions resulting from noise and other stochastic processes such as PMD. However, the optimal nonlinear equaliser for an arbitrary nonlinear channel is known in the soft-decision Viterbi decoder. Although the Viterbi algorithm is too complex to be implemented in hardware for high-level modulation formats in the near-future, its optimality may give a useful upper bound on the potential of nonlinearity compensation and a performance benchmark for current algorithms. This will indicate the usefulness of recent research in both backpropagation algorithm development and nonlinear information-theoretic capacity, which is also a field reliant on backpropagation.

6.1.2 DSP Algorithms for Optimal 4D Modulation Formats

Although our research into DSP algorithms for PS-QPSK have demonstrated that a useful improvement in performance may be obtained by applying information-theoretic results regarding the optical channel, these results are limited to the low spectral-efficiency regime. The pioneering work presented in [34], however

presented many optimal constellations of different orders, only two of which have been investigated to date. Higher order modulation formats which are optimal in the power-constrained case should be investigated to determine if it is possible to obtain significant benefits over the current candidates for high spectral-efficiency regime such as DP-QAM16 and DP-QAM64. This research should encompass bit-mapping, coding schemes and post-processing for equalisation and phase and polarization recovery.

6.1.3 Equaliser Initialisation Algorithms

The equaliser initialisation algorithm presented in chapter 4 represents a novel approach to the mitigation of equaliser singular mal-convergence. Algorithms of this kind offer a great potential to improve robustness and convergence times of current systems, without significant increases in complexity. However, the algorithm presented in chapter 4 has an inherent limitation in that it works only for PS-QPSK modulation. We propose that by extending the general principle of this algorithm it should be possible to mitigate the possibility of singular mal-convergence without the need for independent component analysis (ICA) or other such bulky algorithms.

6.2 Conclusions

In this thesis, we have examined DSP algorithms for coherent optical communication systems. We have focussed on algorithms which offer performance benefits through the compensation of fibre nonlinearity, and algorithms which enable the use of PS-QPSK modulation, which offers an improvement in noise tolerance over conventional dual-polarization formats.

We have investigated the performance of a coherent receiver with nonlinearity-compensating DSP and have shown that it can be successfully used to mitigate intra-channel nonlinearities in both DP-QPSK and DP-QAM16 over distances of 7780 km and 1600 km respectively. The impact of the key receiver DSP parameter, namely the nonlinear step size was investigated. It was shown that significant improvements in performance may be achieved with resolution significantly coarser than a single span. While performance in this *long-step* region may be improved with the use of a three block Wiener-Hammerstein model rather than the more commonly used Wiener model, the increased computational effort this model requires offsets any

benefit when performance is examined in terms of the mean dispersive block length. For the examined receiver bandwidth and symbol rate, the benefit of nonlinearity compensation saturated for a nonlinear step size of 160 km for DP-QPSK and 80 km for DP-QAM16. Additionally, nonlinear backpropagation appears to offer a greater benefit for DP-QAM16, which may be attributed to this format's higher susceptibility to fibre nonlinear effects. This leads us to infer that nonlinearity compensation of this kind is considerably more attractive for modulation formats which are highly spectrally efficient, and transmitted over short links, where the reduced memory due to dispersion and the increased benefit available combine to produce greater benefits from fewer blocks.

We then compared the complexity of nonlinear backpropagation with that of chromatic dispersion compensation only. An approximately exponential relationship was found between complexity in terms of the required number of complex multipliers and performance in Q-factor for both modulation formats. A 1 dBQ improvement in performance requires an increase in complexity of approximately a factor of 10 for DP-QAM16 and a factor of approximately 15 for DP-QPSK. This result indicates that even in the single channel case, the available improvement in performance when using nonlinear backpropagation is limited, while the computational cost is high.

While the benefits of nonlinearity compensation may be considerable in some cases, the computational cost of implementation makes these algorithms prohibitive at the present time. Additionally, the difficulties of performing inter-channel nonlinearity compensation require that information is shared between adjacent WDM channels requires that bussing data to each receiver at a rate in the hundreds of Gb/s is required. Perhaps the most damning of all is that in a ROADM network, not all signals that generate inter-channel nonlinearities may be at every node, as some signals may have already been dropped. While digital nonlinearity compensation is at best a distant possibility, better choice of modulation format may provide better performance than conventional coherent DP-QPSK systems at 40G and 100G line rates.

Power efficient polarization-switched QPSK (PS-QPSK) modulation was then examined as a low-complexity means of achieving performance superior to

conventional systems in the low spectral efficiency regime. We examined the efficiency of several bit mapping schemes for PS-QPSK and used them to characterise the performance of novel DSP algorithms for this modulation format.

A novel polarization-switched CMA equaliser was proposed and was found to have comparable performance to the DP-CMA equaliser, without suffering from degenerate mal-convergence with certain input polarizations states or high levels of PDL. The PS-CMA may therefore be said to be more robust than the DP-CMA without the use of training sequences or significant additional algorithmic complexity.

A comparison was performed between PS-QPSK and DP-QPSK for 100 GbE WDM systems over a 50 GHz frequency grid. PS-QPSK was found to enable a gain in receiver sensitivity of 1 dB in required OSNR when compared with standard DP-QPSK and a 7% FEC overhead.

We then presented the first experimental measurements of the transmission performance of PS-QPSK at 42.9 Gb/s using a simple and easily realisable generation technique. The implementation penalty for PS-QPSK modulation was found to be less than 1 dB at a BER of 3.8×10^{-3} . This transmitter was then used to perform a characterisation of transmission performance for PS-QPSK, comparing PS-QPSK and DP-QPSK at 42.9 Gb/s over a 50 GHz frequency grid. An improvement in launch power margin (the range of launch powers for which the FEC limit may be maintained) of greater than 1 dB was found in all cases. The optimum launch power for both modulation formats was found to be -3.5 dBm per channel, while maximum reach was increased by more than 30% from 10350 km to 13640 km. At the time of publication, these results represented the longest distance 40 Gb/s WDM transmission achieved over an uncompensated link, with standard fibre and amplification. While standards have been set for 40G and 100G coherent optical systems, it is conceivable that this increased margin from PS-QPSK may be of some practical use in some future systems, for example: access networks; submarine systems or as a fall-back modulation format for software defined optical transceivers.

It is clear that by tailoring modulation and digital algorithms to the optical channel, improvements in performance may be obtained over standard algorithms and

systems. While the computational complexity of digital nonlinearity compensation is currently prohibitive, with the scaling of CMOS technology, it is conceivable that some form of digital nonlinearity compensation will be implemented in future systems. As we look to future systems, it also seems likely that modulation formats such as PS-QPSK which are better suited to the optical channel will be used, as coherent detection becomes ubiquitous and better developed.

7 References

- [1] K. C. Kao and G. a Hockham, "Dielectric-fibre surface waveguides for optical frequencies," *IEE Proceedings J Optoelectronics*, vol. 133, no. 3, p. 191, 1986.
- [2] R. S. Vodhanel, A. F. Elrefaie, M. Z. Iqbal, R. E. Wagner, J. L. Gimlett, and S. Tsuji, "Performance of directly modulated DFB lasers in 10-Gb/s ASK, FSK, and DPSK lightwave systems," *Lightwave Technology, Journal of*, vol. 8, no. 9, pp. 1379-1386, 1990.
- [3] K. Kikuchi, T. Okoshi, and J. Kitano, "Measurement of bit-error rate of heterodyne-type optical communication system—A simulation experiment," *Quantum Electronics, IEEE Journal of*, vol. 17, no. 12, pp. 2266–2267, 1981.
- [4] R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μ m," *Electronics Letters*, vol. 23, no. 19, p. 1026, 1987.
- [5] C. A. Brackett, "Dense wavelength division multiplexing networks: principles and applications," *Selected Areas in Communications, IEEE Journal on*, vol. 8, no. 6, pp. 948-964, 1990.
- [6] P. J. Winzer and R. J. Essiambre, "Advanced optical modulation formats," *Proceedings of the IEEE*, vol. 94, no. 5, pp. 952–985, 2006.
- [7] J. M. Kahn and K.-P. Ho, "Spectral Efficiency Limits and Modulation/Detection Techniques for DWDM Systems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 10, no. 2, pp. 259-272, Mar. 2004.
- [8] H. Sun, K. T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Optics Express*, vol. 16, no. 2, pp. 873–879, 2008.
- [9] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Optics Express*, vol. 16, no. 2, pp. 753–791, 2008.
- [10] S. J. Savory, "Digital filters for coherent optical receivers," *Optics Express*, vol. 16, no. 2, pp. 804–817, 2008.
- [11] "Implementation Agreement for Integrated Dual Polarization Intradyn Coherent Receivers," *Optical Internetworking Forum*, 2010. [Online]. Available: www.oiforum.com/public/documents/OIF_DPC_RX-01.0.pdf.
- [12] D. S. Millar et al., "Mitigation of Fiber Nonlinearity Using a Digital Coherent Receiver," *Selected Topics in Quantum Electronics, Journal of*, vol. 16, no. 5, pp. 1217–1226, 2010.

- [13] D. S. Millar and S. J. Savory, "Blind adaptive equalization of polarization-switched QPSK modulation," *Optics Express*, vol. 19, no. 9, pp. 8533–8538, 2011.
- [14] P. Johannisson, M. Sjödin, M. Karlsson, H. Wymeersch, E. Agrell, and P. A. Andrekson, "Modified constant modulus algorithm for polarization-switched QPSK," *Optics Express*, vol. 19, no. 8, pp. 7734-7741, 2011.
- [15] D. S. Millar et al., "Generation and long-haul transmission of polarization-switched QPSK at 42.9 Gb/s," *Optics Express*, vol. 19, no. 10, pp. 9296-9302, 2011.
- [16] A. Kanda, A. Ohki, Y. Suzuki, and Y. Akatsu, "10 Gbit/s small form factor optical transceiver for 40 km WDM transmission," *Electronics Letters*, vol. 40, no. 8, pp. 494–495, 2004.
- [17] M. Nakazawa, S. Okamoto, T. Omiya, K. Kasai, and M. Yoshida, "256 QAM (64 Gbit/s) coherent optical transmission over 160 km with an optical bandwidth of 5.4 GHz," in *Optical Fiber Communication Conference*, 2010, p. OMJ5.
- [18] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *Lightwave Technology, Journal of*, vol. 25, no. 8, pp. 2033–2043, 2007.
- [19] R. A. Linke and A. H. Gnauck, "High-capacity coherent lightwave systems," *Lightwave Technology, Journal of*, vol. 6, no. 11, pp. 1750-1769, 1988.
- [20] T. Okoshi, "Heterodyne and Coherent Optical Fiber Communications: Recent Progress," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 30, no. 8, pp. 1138-1149, Aug. 1982.
- [21] B. Glance, K. Pollock, C. Burrus, B. Kasper, G. Eisenstein, and L. Stulz, "Densely spaced WDM coherent optical star network," *Electronics Letters*, vol. 23, no. 17, pp. 875–876, 1987.
- [22] D. Breuer et al., "Measurements of PMD in the installed fiber plant of Deutsche Telekom," in *Digest of the LEOS Summer Topical Meetings, 2003*, 2003, vol. 1, pp. MB2–1.
- [23] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver," *Optics Express*, vol. 15, no. 5, pp. 2120–2126, 2007.
- [24] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *Selected Topics in Quantum Electronics, Journal of*, vol. 16, no. 5, pp. 1164–1179, 2010.

- [25] L. Anderson, "The PIN junction photodiode as a detector of light modulated at microwave frequencies," *Solid-State Circuits Conference. Digest of*, p. FPM12.4, 1963.
- [26] D. S. Millar, S. Makovejs, V. Mikhailov, R. Killely, P. Bayvel, and S. Savory, "Experimental comparison of nonlinear compensation in long-haul PDM-QPSK transmission at 42.7 and 85.4 Gb/s," in *European Conference on Optical Communication*, 2009, p. 9.4.4.
- [27] F. Derr, "Coherent Optical QPSK Intradynic System: Concept and Digital Receiver Realization," *Lightwave Technology, Journal of*, vol. 10, no. 9, pp. 1290-1296, 1992.
- [28] C. Fludger and T. Duthel, "Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission," *Lightwave Technology, Journal of*, vol. 26, no. 1, pp. 64-72, 2008.
- [29] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communications Systems*. Plenum Press, 1992, pp. 271-272.
- [30] C. Cahn, "Combined Digital Phase and Amplitude Modulation Communication Systems," *Communications, IRE Transactions on*, vol. 8, no. 3, pp. 150-155, Sep. 1960.
- [31] P. M. Krummrich, E.-D. Schmidt, W. Weiershausen, and A. Mattheus, "Field trial results on statistics of fast polarization changes in long haul WDM transmission systems," in *Optical Fiber Communication Conference*, 2005, p. 3 pp. Vol. 4.
- [32] U. V. Koc, "PLL-free quadrature-amplitude modulation in coherent optical communication," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, 2007, pp. 2299-2302.
- [33] W. Shieh, H. Bao, and Y. Tang, "Coherent optical OFDM: theory and design," *Optics Express*, vol. 16, no. 2, pp. 841-859, 2008.
- [34] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *Lightwave Technology, Journal of*, vol. 27, no. 22, pp. 5115-5126, 2009.
- [35] M. Proakis, J. G. Salehi, *Digital Communications*, 5th ed. McGraw-Hill, 2008.
- [36] E. Ip and J. M. Kahn, "Carrier synchronization for 3-and 4-bit-per-symbol optical transmission," *Lightwave Technology, Journal of*, vol. 23, no. 12, pp. 4110-4124, 2005.
- [37] X. Zhou et al., "Transmission of 32-Tb/s Capacity Over 580 km Using RZ-Shaped PDM-8QAM Modulation Format and Cascaded Multimodulus Blind

- Equalization Algorithm,” *Lightwave Technology, Journal of*, vol. 28, no. 4, pp. 456-465, 2010.
- [38] M. Seimetz, M. Noelle, and E. Patzak, “Optical systems with high-order DPSK and star QAM modulation based on interferometric direct detection,” *Lightwave Technology, Journal of*, vol. 25, no. 6, pp. 1515–1530, 2007.
- [39] E. L. Wooten et al., “A review of lithium niobate modulators for fiber-optic communications systems,” *Selected Topics in Quantum Electronics, Journal of*, vol. 6, no. 1, pp. 69-82, 2000.
- [40] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. Academic Press, 2001, pp. 45-51.
- [41] J. P. Gordon and H. Kogelnik, “PMD fundamentals: Polarization mode dispersion in optical fibers,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 9, p. 4541, 2000.
- [42] T. Okoshi, K. Emura, K. Kikuchi, and R. T. Kersten, “Computation of Bit-Error Rate of Various Heterodyne and Coherent-Type Optical Communication Schemes,” *Journal of Optical Communications*, vol. 2, no. 3, pp. 89-96, Sep. 1981.
- [43] M. G. Taylor, “Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments,” *Photonics Technology Letters*, vol. 16, no. 2, pp. 674–676, 2004.
- [44] R. Noe, “PLL-free synchronous QPSK polarization multiplex/diversity receiver concept with digital I&Q baseband processing,” *Photonics Technology Letters*, vol. 17, no. 4, pp. 887–889, 2005.
- [45] F. M. Gardner, “Interpolation in Digital Modems-Part I: Fundamentals,” *Communications, IEEE Transactions on*, vol. 41, no. 3, pp. 501-507, 1993.
- [46] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 2001.
- [47] Y. Sato, “A Method of Self-Recovering Equalization for Multilevel Amplitude-Modulation Systems,” *Communications, IEEE Transactions on*, vol. 23, no. 6, pp. 679-682, Jun. 1975.
- [48] B. Widrow and M. E. Hoff, “Adaptive switching circuits.,” *Proc. IRE WESCON*, pp. 96-104, 1960.
- [49] G. Picchi and G. Prati, “Blind equalization and carrier recovery using a stop-and-go decision-directed algorithm,” *Communications, IEEE Transactions on*, vol. 35, no. 9, pp. 877–887, 1987.
- [50] D. Godard, “Self-recovering equalization and carrier tracking in two-dimensional data communication systems,” *Communications, IEEE Transactions on*, vol. 28, no. 11, pp. 1867–1875, 1980.

- [51] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *Lightwave Technology, Journal of*, vol. 27, no. 15, pp. 3042–3049, 2009.
- [52] T. Pfau, N. Kaneda, S. Corteselli, A. Leven, and Y. K. Chen, "Real-Time FPGA-Based Intradynic Coherent Receiver for 40 Gbit/s Polarization-Multiplexed 16-QAM," in *Optical Fiber Communication Conference*, 2011, p. OTuN4.
- [53] D. Rife, "Single tone parameter estimation from discrete-time observations," *Information Theory, IEEE Transactions*, 1974.
- [54] G. De Jonghe, "Asymptotic expression for the equivocation probability of the NDA FF carrier synchroniser," *Electronics Letters*, vol. 29, no. 24, pp. 2077–2078, 1993.
- [55] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *Information Theory, IEEE Transactions on*, vol. 29, no. 4, pp. 543–551, 1983.
- [56] F. Rice and M. Rice, "A new algorithm for 16QAM carrier phase estimation using QPSK partitioning," *Digital Signal Processing*, vol. 86, no. 2002, pp. 77-86, 2002.
- [57] Y. Gao, A. Lau, C. Lu, J. Wu, and Y. Li, "Low-Complexity Two-Stage Carrier Phase Estimation for 16-QAM Systems using QPSK Partitioning and Maximum Likelihood Detection," in *Optical Fiber Communication Conference*, 2011, vol. 2, p. OMJ6.
- [58] S. Makovejs et al., "Characterization of long-haul 112Gbit/s PDM-QAM-16 transmission with and without digital nonlinearity compensation," *Optics express*, vol. 18, no. 12, pp. 12939-47, Jun. 2010.
- [59] A. H. Gnauck, P. J. Winzer, S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, "Spectrally Efficient Long-Haul WDM Transmission Using 224-Gb/s Polarization-Multiplexed 16-QAM," *Lightwave Technology, Journal of*, vol. 29, no. 4, pp. 373-377, Feb. 2011.
- [60] S. Makovejs et al., "Novel method of generating QAM-16 signals at 21.3 Gbaud and transmission over 480 km," *Photonics Technology Letters*, vol. 22, no. 1, pp. 36–38, 2010.
- [61] S. Makovejs et al., "Experimental investigation of PDM-QAM16 transmission at 112 Gbit/s over 2400 km," in *Optical Fiber Communication Conference*, 2010, p. OMJ6.
- [62] X. Li, X. Chen, G. Goldfarb, E. Mateo, and I. Kim, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Optics Express*, vol. 16, no. 2, pp. 880-888, 2008.

- [63] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear effects using digital backpropagation," *Lightwave Technology, Journal of*, vol. 26, no. 20, pp. 3416-3425, 2008.
- [64] G. Charlet, M. Salsi, P. Tran, M. Bertolini, H. Mardoyan, and J. Renaudier, "72x100Gb/s transmission over transoceanic distance, using large effective area fiber, hybrid Raman-Erbium amplification and coherent detection," *Optical Fiber Communication Conference*, p. PDPB6, 2009.
- [65] G. Goldfarb, M. G. Taylor, and G. Li, "Experimental demonstration of fiber impairment compensation using the split-step finite-impulse-response filtering method," *Photonics Technology Letters*, vol. 20, no. 22, pp. 1887–1889, 2008.
- [66] D. Marcuse, C. Menyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *Lightwave Technology, Journal of*, vol. 15, no. 9, pp. 1735–1746, 1997.
- [67] C. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber transmission," *Journal of engineering mathematics*, vol. 36, no. 1, pp. 113–136, 1999.
- [68] S. Oda et al., "112 Gb/s DP-QPSK transmission using a novel nonlinear compensator in digital coherent receiver," in *Optical Fiber Communication Conference*, 2009, vol. 1, no. c, p. OThR6.
- [69] F. Yaman and G. Li, "Nonlinear impairment compensation for polarization-division multiplexed WDM transmission using digital backward propagation," *Photonics Journal, IEEE*, vol. 1, no. 2, pp. 144–152, 2009.
- [70] G. H. Weiss and A. A. Maradudin, "The Baker-Hausdorff Formula and a Problem in Crystal Physics," *Journal of Mathematical Physics*, vol. 3, no. 4, pp. 771-777, 1962.
- [71] R. Haber and H. Unbehauen, "Structure identification of nonlinear dynamic systems--A survey on input/output approaches," *Automatica*, vol. 26, no. 4, pp. 651-677, 1990.
- [72] E. Yamazaki et al., "Multi-staged nonlinear compensation in coherent receiver for 16,340-km transmission of 111-Gb/s no-guard-interval co-OFDM," in *European Conference on Optical Communication*, 2009, p. 9.4.6.
- [73] T. Tanimura et al., "Systematic analysis on multi-segment dual-polarisation nonlinear compensation in 112 Gb/s DP-QPSK coherent receiver," in *European Conference on Optical Communication*, 2009, no. 1, p. 9.4.5.
- [74] L. Du, B. Schmidt, and A. Lowery, "Efficient digital backpropagation for PDM-CO-OFDM optical transmission systems," in *Optical Fiber Communication Conference*, 2010, p. OTuE2.

- [75] S. J. Savory, G. Gavioli, E. Torrenco, and P. Poggiolini, "Impact of interchannel nonlinearities on a split-step intrachannel nonlinear equalizer," *Photonics Technology Letters*, vol. 22, no. 10, pp. 673–675, 2010.
- [76] X. Liu, S. Chandrasekhar, B. Zhu, P. J. Winzer, and D. W. Peckham, "7 x 224-Gb/s WDM Transmission of Reduced-Guard-Interval CO-OFDM with 16-QAM Subcarrier Modulation on a 50-GHz Grid over 2000 km of ULAF and Five ROADMs Passes," in *European Conference on Optical Communication*, 2010, no. c, p. Tu.3.C.2.
- [77] X. Liu, S. Chandrasekhar, B. Zhu, P. J. Winzer, A. H. Gnauck, and D. W. Peckham, "448-Gb/s reduced-guard-interval CO-OFDM transmission over 2000 km of ultra-large-area fiber and five 80-GHz-grid ROADMs," *Lightwave Technology, Journal of*, vol. 29, no. 4, pp. 483–490, 2011.
- [78] S. Makovejs, E. Torrenco, D. S. Millar, R. I. Killely, S. J. Savory, and P. Bayvel, "Comparison of Pulse Shapes in a 224Gbit/s (28Gbaud) PDM-QAM16 Long-Haul Transmission Experiment," in *Optical Fiber Communication Conference*, 2011, p. OMR5.
- [79] L. Li et al., "Implementation efficient nonlinear equalizer based on correlated digital backpropagation," in *Optical Fiber Communication Conference*, 2011, no. 2, p. OWW3.
- [80] T. Tanimura, S. Oda, T. Hoshida, L. Li, Z. Tao, and J. C. Rasmussen, "Experimental Characterization of Nonlinearity Mitigation by Digital Back Propagation and Nonlinear Polarization Crosstalk Canceller under High PMD condition," in *Optical Fiber Communication Conference*, 2011, vol. 1, no. c, p. JWA20.
- [81] E. Yamazaki, A. Sano, T. Kobayashi, E. Yoshida, and Y. Miyamoto, "Mitigation of Nonlinearities in Optical Transmission Systems," in *Optical Fiber Communication Conference*, 2011, p. OThF1.
- [82] L. Zhu, F. Yaman, and G. Li, "Experimental demonstration of XPM compensation for WDM fibre transmission," *Electronics letters*, vol. 46, no. 16, pp. 1140–1141, 2010.
- [83] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *Information Theory, IEEE Transactions on*, vol. 13, no. 2, pp. 260–269, 1967.
- [84] A. Gorshtein, D. Sadot, G. Katz, and O. Levy, "Coherent CD Equalization for 111Gbps DP-QPSK with One Sample per Symbol Based on Anti-Aliasing Filtering and MLSE," in *Optical Fiber Communication Conference*, 2010, p. OThT2.
- [85] P. Poggiolini, "MLSE receivers: Application scenarios, fundamental limits and experimental validations," in *European Conference on Optical Communication*, 2008, vol. 2, no. September, p. Tu.1.D.1.

- [86] J. Zhao and A. Ellis, "Performance improvement using a novel MAP detector in coherent WDM systems," in *European Conference on Optical Communication*, 2008, vol. 2, no. September, p. Tu.1.D.2.
- [87] Y. Cai et al., "Experimental demonstration of coherent MAP detection for nonlinearity mitigation in long-haul transmissions," in *Optical Fiber Communication Conference*, 2010, p. OTuE1.
- [88] K. V. Peddanarappagari and M. Brandt-Pearce, "Volterra series approach for optimizing fiber-optic communications system designs," *Lightwave Technology, Journal of*, vol. 16, no. 11, pp. 2046–2055, 1998.
- [89] B. Xu and M. Brandt-Pearce, "Modified Volterra series transfer function method," *Photonics Technology Letters*, vol. 14, no. 1, pp. 47–49, 2002.
- [90] R. Weidenfeld, M. Nazarathy, R. Noe, and I. Shpantzer, "Volterra nonlinear compensation of 112 Gb/s ultra-long-haul coherent optical OFDM based on frequency-shaped decision feedback," in *European Conference on Optical Communication*, 2009, no. 1, p. 2.3.3.
- [91] R. Weidenfeld, M. Nazarathy, R. Noe, and I. Shpantzer, "Volterra nonlinear compensation of 100G coherent OFDM with baud-rate ADC, tolerable complexity and low intra-channel FWM/XPM error propagation," in *Optical Fiber Communication Conference*, 2010, no. iii, p. OTuE3.
- [92] Z. Pan, B. Châtelain, M. Chagnon, and D. V. Plant, "Volterra Filtering for Nonlinearity Impairment Mitigation in DP-16QAM and DP-QPSK Fiber Optic Communication Systems," in *Optical Fiber Communication Conference*, 2011, no. 3, p. JThA40.
- [93] K. Roberts, C. Li, L. Strawczynski, M. O'Sullivan, and I. Hardcastle, "Electronic precompensation of optical nonlinearity," *Photonics Technology Letters*, vol. 18, no. 2, pp. 403–405, 2006.
- [94] L. Du and A. Lowery, "Improving nonlinear precompensation in direct-detection optical OFDM communications systems," in *European Conference on Optical Communication*, 2008, vol. 5, no. September, p. P.4.08.
- [95] J. K. Fischer and K. Petermann, "Scaling of nonlinear threshold in WDM transmission systems using electronic precompensation of intrachannel nonlinearities," in *European Conference on Optical Communication*, 2009, p. P4.05.
- [96] P. M. Watts et al., "10 . 7 Gb / s Electronically Predistorted Transmission over 800 km Standard Single Mode Fibre using FPGA-based Real-Time Processing," *Electronics Letters*, vol. 3, no. September, pp. 121-122, 2008.
- [97] A. Chraplyvy, "The coming capacity crunch," *European Conference on Optical Communication*, p. Mo1.0.2, 2009.

- [98] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379-423,623-656, 1948.
- [99] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Lightwave Technology, Journal of*, vol. 28, no. 4, pp. 662–701, 2010.
- [100] P. Winzer, A. Gnauck, C. Doerr, M. Magarini, and L. Buhl, "Spectrally efficient long-haul optical networking using 112-Gb/s polarization-multiplexed 16-QAM," *Lightwave Technology, Journal of*, vol. 28, no. 4, pp. 547–556, 2010.
- [101] C. Xie, "Suppression of inter-channel nonlinearities in WDM coherent PDM-QPSK systems using periodic-group-delay dispersion compensators," in *European Conference on Optical Communication*, 2009, p. P4.08.
- [102] E. F. Mateo, F. Yaman, and G. Li, "Efficient compensation of inter-channel nonlinear effects via digital backward propagation in WDM optical transmission," *Optics Express*, vol. 18, no. 14, pp. 15144–15154, 2010.
- [103] S. Makovejs et al., "Experimental investigation of PDM-QAM16 transmission at 112 Gbit/s over 2400 km," in *Optical Fiber Communication Conference*, 2010, pp. 16-18.
- [104] A. H. Gnauck, R. Tkach, A. Chraplyvy, and T. Li, "High-capacity optical transmission systems," *Lightwave Technology, Journal of*, vol. 26, no. 9, pp. 1032–1045, 2008.
- [105] S. Makovejs, G. Gavioli, V. Mikhailov, R. I. Killey, and P. Bayvel, "Experimental and numerical investigation of bit-wise phase-control OTDM transmission," *Optics express*, vol. 16, no. 23, pp. 18725–18730, 2008.
- [106] S. J. Savory, "Optimum electronic dispersion compensation strategies for nonlinear transmission," *Electronics Letters*, vol. 42, no. 7, pp. 2-3, 2006.
- [107] J. C. Geyer, C. R. Fludger, T. Duthel, C. Schuijen, and B. Schmauss, "Efficient Frequency Domain Chromatic Dispersion Compensation in a Coherent Polmux QPSK-Receiver," in *Optical Fiber Communication Conference*, 2010, vol. 2, p. OWV5.
- [108] C. R. Fludger, J. C. Geyer, and T. Duthel, "Digital Signal Processing for the Realisation of 40 and 100G CP-QPSK Transponders," *Signal Processing in Photonic Communications*, p. SPTuC1, 2010.
- [109] Akihide Sano et al., "69.1-Tb/s (432 x 171-Gb/s) C- and Extended L-Band Transmission over 240 Km Using PDM-16-QAM Modulation and Digital Coherent Detection - OSA Technical Digest (CD)," in *Optical Fiber Communication Conference*, 2010, p. PDPB7.

- [110] L. Zhu, X. Li, E. Mateo, and G. Li, “Complementary FIR Filter Pair for Distributed Impairment Compensation of WDM Fiber Transmission,” *Photonics Technology Letters*, vol. 21, no. 5, pp. 292–294, 2009.
- [111] M. Karlsson and E. Agrell, “Which is the most power-efficient modulation format in optical links?,” *Optics Express*, vol. 17, no. 13, pp. 10814–10819, 2009.
- [112] H. Bülow, “Polarization QAM modulation (POL-QAM) for coherent detection schemes,” *Optical Fiber Communication Conference*, no. 1, p. OWG2, 2009.
- [113] “40G Submarine Applications,” *Ciena Corp. Application Note*. [Online]. Available: www.ciena.com.
- [114] “Alcatel-Lucent 1626 Light Manager, Data Sheet release 6.1,” *Alcatel-Lucent*. [Online]. Available: www.alcatel-lucent.com.
- [115] L. E. Nelson, X. Zhou, N. M. Suibhne, A. D. Ellis, and P. Magill, “Experimental comparison of coherent polarization-switched QPSK to polarization- multiplexed QPSK for 10×100 km WDM transmission,” *Optics Express*, vol. 19, no. 11, pp. 10849-10856, 2011.
- [116] P. Poggiolini, G. Bosco, A. Carena, V. Curri, and F. Forghieri, “Performance evaluation of coherent WDM PS-QPSK (HEXA) accounting for non-linear fiber propagation effects,” *Optics Express*, vol. 18, no. 11, pp. 11360–11371, 2010.
- [117] P. Serena, A. Vannucci, and A. Bononi, “The performance of polarization switched-QPSK (PS-QPSK) in dispersion managed WDM transmissions,” in *European Conference on Optical Communication*, 2010, p. Th.10.E.2.
- [118] G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, “Robust multilevel coherent optical systems with linear processing at the receiver,” *Lightwave Technology, Journal of*, vol. 27, no. 13, pp. 2357–2369, 2009.
- [119] I. Fatadin, D. Ives, and S. J. Savory, “Compensation of frequency offset for differentially encoded 16-and 64-QAM in the presence of laser phase noise,” *Photonics Technology Letters*, vol. 22, no. 3, pp. 176–178, 2010.
- [120] N. S. Bergano, F. Kerfoot, and C. Davidsion, “Margin measurements in optical amplifier system,” *Photonics Technology Letters*, vol. 5, no. 3, pp. 304–306, 1993.
- [121] M. Kushnerov et al., “DSP for coherent single-carrier receivers,” *Lightwave Technology, Journal of*, vol. 27, no. 16, pp. 3614–3622, 2009.
- [122] A. Vgenis, C. S. Petrou, C. B. Papadias, I. Roudas, and L. Raptis, “Nonsingular constant modulus equalizer for PDM-QPSK coherent optical receivers,” *Photonics Technology Letters*, vol. 22, no. 1, pp. 45–47, 2010.

- [123] J. X. Cai et al., "112x112 Gb/s Transmission over 9,360 km with Channel Spacing Set to the Baud Rate (360% Spectral Efficiency)," in *European Conference on Optical Communication*, 2010, p. PD2.1.
- [124] A. Sano et al., "100x120-Gb/s PDM-64-QAM Transmission over 160 km Using Linewidth-Tolerant Pilotless Digital Coherent Detection," in *European Conference on Optical Communication*, 2010, p. PD2.4.
- [125] M. Sjödin, P. Johannisson, H. Wymeersch, P. A. Andrekson, and M. Karlsson, "Comparison of polarization-switched QPSK and polarization-multiplexed QPSK at 30 Gbit / s," *Optics Express*, vol. 19, no. 8, pp. 7839-7846, 2011.
- [126] E. Torrenco et al., "Influence of Pulse Shape in 112-Gb/s WDM PDM-QPSK Transmission," *Photonics Technology Letters*, vol. 22, no. 23, pp. 1714-1716, 2010.