“Investigation of Degrading Effects and Performance Optimization in Long-Haul WDM Transmission Systems and Reconfigurable Networks”

by

Lianshan Yan

CSI-04-11-01
Investigation of Degrading Effects and Performance Optimization in Long-Haul WDM Transmission Systems and Reconfigurable Networks

by

Lianshan Yan

A Dissertation Presented to the
FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
in Partial Fulfillment of the
Requirement for the Degree
DOCTOR OF PHILOSOPHY
(Electrical Engineering)

Committee members
Alan E. Willner (chairman)
William H. Steier
Mark Thompson

November 2004

Copyright 2004

Lianshan Yan
To my beloved parents, sister and tianhong for their everlasting love and support.
Acknowledgements

I would like to thank my advisor and dissertation committee chairman, Professor Alan E. Willner, for his invaluable attention, guidance, and support throughout my graduate studies. I would also like to extend my great appreciation to the other dissertation committee members, Professor William H. Steier and Professor Mark Thompson. In addition, I would like to thank Dr. X. Steve Yao from General Photonics Corp., Dr. Y. Shi from Boeing Satellite Corp., Prof. W. L. Kath from Northwestern University, Professor C. R. Menyuk from University of Maryland at Baltimore Country, Prof. D. Yevick from University of Waterloo, Dr. Q. Yu from Scientera Networks and Dr. Y. Xie from Intel, for their particular influence on my way of thinking and working. I would also thank the other collaborators such as Dr. Ivan T. Lima from UMBC (now at North Dakota State University), Dr. Y. Chang from Vitesse Semiconductor Inc., Dr. L. Paraschis from Cisco Systems Inc., T. Lu from University of Waterloo.

I would like to pay my sincere thanks to my colleagues from Optical Communication Laboratory (OCLAB) for their generous help throughout my graduate career. And finally I would like to thank those people who have contributed in some way to the success of my academic endeavors. They are Milly Montenegro, Mayumi Thrasher, and Gerrielyn Ramos.
# Table of contents

1 INTRODUCTION................................................................................................................. 1

2 BACKGROUND ...................................................................................................................... 4
   2.1 FIBER DISPERSION................................................................................................. 4
   2.2 FIBER NONLINEARITIES.......................................................................................... 10
      2.2.1 Self-Phase Modulation................................................................................... 11
      2.2.2 Cross-Phase Modulation................................................................................. 13
      2.2.3 Four-Wave-Mixing....................................................................................... 15
      2.2.4 Stimulated Scattering.................................................................................... 17
   2.3 POLARIZATION EFFECTS ....................................................................................... 18
      2.3.1 Background of Polarization.......................................................................... 18
      2.3.2 Polarization-Mode-Dispersion (PMD)........................................................... 25
      2.3.3 Polarization-Dependent-Loss (PDL)............................................................... 30
   2.4 RECIRCULATING LOOP TESTBED ............................................................................. 37
   2.5 SUMMARY................................................................................................................. 40

3 INVESTIGATION OF DEGRADING EFFECTS ............................................................. 41
   3.1 LOOP-SYNCHRONOUS POLARIZATION SCRAMBLING TECHNIQUE FOR SIMULATING POLARIZATION EFFECTS USING RECIRCULATING FIBER LOOPS ......................................................... 41
   3.2 COMBINED EFFECTS OF PMD AND PDL IN LONG-HAUL SYSTEMS .................... 55
   3.3 DELETERIOUS SYSTEM EFFECTS DUE TO LOW FREQUENCY POLARIZATION SCRAMBLING IN THE PRESENCE OF PDL .................................................................................................................... 63
   3.4 SYSTEM IMPACTS BY THE GROUP-DELAY RIPPLE IN SINGLE AND CASCADED CHIRPED FIBER BRAGG GRATINGS ......................................................................................................................... 70

4 TECHNICAL SOULITONS FOR PERFORMANCE OPTIMIZATION IN WDM SYSTEMS AND RECONFIGURABLE NETWORKS .............................................................. 80
   4.1 REACH EXTENSION IN 10-Gb/s AND 40-Gb/s DIRECTLY MODULATED TRANSMISSION SYSTEMS USING ASYMMETRIC NARROWBAND OPTICAL FILTERING ................................................................. 80
   4.2 MEASUREMENT OF THE CHIRP PARAMETER OF ELECTRO-OPTIC MODULATORS BY COMPARING THE PHASE BETWEEN TWO SIDE BANDS ......................................................................................... 105
   4.3 CHARACTERIZATION OF THE CHIRP PARAMETER OF ELECTRO-OPTIC INTENSITY MODULATORS USING OPTICAL SPECTRUM ANALYSIS .............................................................................................. 114
   4.4 PERFORMANCE OPTIMIZATION OF RZ DATA FORMAT IN WDM SYSTEMS USING TUNABLE PULSE-WIDTH MANAGEMENT AT THE TRANSMITTER ...................................................................................... 121
   4.5 UNIFORMLY DISTRIBUTED STATES OF POLARIZATION ON THE POINCARÈ SPHERE USING AN IMPROVED POLARIZATION SCRAMBLING SCHEME ................................................................. 131
   4.6 DEMONSTRATION OF IN-LINE MONITORING AND COMPENSATION OF POLARIZATION DEPENDENT LOSS FOR MULTIPLE CHANNELS ......................................................................................... 146
   4.7 PROGRAMMABLE GROUP DELAY MODULE USING BINARY POLARIZATION SWITCHING .... 156
   4.8 POLARIZATION-MODE-DISPERSION EMULATOR USING VARIABLE DIFFERENTIAL-GROUP-DELAY (DGD) ELEMENTS AND ITS USE FOR EXPERIMENTAL IMPORTANCE SAMPLING ............. 180
   4.9 DIFFERENTIAL GROUP DELAY MONITORING USED AS FEED FORWARD INFORMATION FOR POLARIZATION MODE DISPERSION COMPENSATION ................................................................................. 201
   4.10 SIMULTANEOUS MONITORING OF BOTH OSNR AND PMD USING POLARIZATION TECHNIQUES ........................................................................................................................................ 208
4.11 Compensation of Higher-Order PMD Using Phase Modulation and Polarization Control in the Transmitter .................................................................................................. 215

4.12 Reach Extension in 10-Gb/s Long-Haul Fiber Links with Adaptive Eye Mapping in a Si-CMOS 16-bit Transceiver IC .................................................................................. 224

5 REFERENCES ........................................................................................................................................................................... 233
List of Tables

Table 1......................................................................................................................29
Table 2......................................................................................................................87
Table 3....................................................................................................................179
List of Figures

Figure 1 Dispersion parameter D versus wavelength of SMF and DSF ...............5
Figure 2 (a) Chromatic dispersion induces optical pulse broadening, proportional to the transmission distance and with the data rate (bits/s). (b) Temperature sensitivity of chromatic dispersion. .........................................................9
Figure 3 Sample eye pattern at different chromatic dispersion value for 10-Gb/s system ................................................................................................................9
Figure 4 The representation of polarization states using the Poincaré sphere ....24
Figure 5 The evolution of the state of polarization as light passes through a linear birefringent plate. .................................................................................................24
Figure 6 Origin of PMD ......................................................................................25
Figure 7 Distribution of first-order PMD (DGD) .................................................27
Figure 8 Graphical representation of the effect of PMD on an optical pulse .......28
Figure 9 Typical eye patterns at various levels of DGD .....................................29
Figure 10 Block diagram of recirculating loop transmission, showing the transmitter/receiver pair, transmitter/loop switches, amplifier chain and the error detection circuitry ....................................................................................38
Figure 11 Optical switch states for (a) load state and (b) loop state ....................38
Figure 12 Experimental setup of a recirculating fiber loop ..................................43
Figure 13 Oscilloscope traces illustrating the state-of-polarization (SOP) of the loop output signal. (a) Without any polarization scrambling inside the loop, a constant output SOP can be obtained by manually adjusting a polarization controller (PC) inside the loop. (b) The SOP is randomized using the loop-synchronous polarization controller, which repeatedly generates a series of random transmission matrixes during every loading-and-circulating cycle. We used a simplified polarimeter consisting of a polarizer followed by an O/E converter of 100 kHz bandwidth to measure an arbitrary component of the SOP vector. The oscilloscope displays the output photocurrent. The gating signal for burst-mode BER measurement is also shown ......................45
Figure 14 Polarization transfer condition of the recirculating loop shown in Figure 12 ..................................................................................................................46
Figure 15 Histogram of received optical powers at $10^{-9}$ BER after circulating 8 times through the loop that contains one element of 8.3-ps DGD. (a) With loop-synchronous polarization scrambling, (b) without polarization scrambling ..................................................................................48
Figure 16 Histogram similar to Figure 15 (a) except that DGD/loop is 5.4 ps ......49
Figure 17 PMD induced power penalties at $10^{-9}$ BER versus the average DGD values .............................................................................................................49
Figure 18 BER curves for (i) back-to-back performance of the PIN receiver used in our experiment, (ii) dispersion managed 772-km transmission with
negligible PMD, (iii) transmission over 772-km fiber of negligible PMD followed by a PM fiber of 43-ps DGD (worst-case polarization splitting) ....50

Figure 19 Histogram of link DGD after circulating 8 times through a loop that contains one fixed DGD element. (a) With loop-synchronous polarization scrambling, (b) without polarization scrambling .............................................52

Figure 20 Schematic diagram of a recirculating loop using multiple loop-synchronous polarizations controllers ..............................................................53

Figure 21 (a) Histogram of link DGD after circulating 8 times the loop shown in Figure 13, with $\Delta \tau_{01} = 1.5$ ps, $\Delta \tau_{02} = 2.5$ ps, and $\Delta \tau_{03} = 3.5$ ps. The inset plots the tail of the cumulative distribution function (CDF) in log scale. The exact Maxwellian distribution is plotted in smooth lines. (b) Autocorrelation function of the PMD vector .............................................................................54

Figure 22 Experimental setup of the recirculating loop testbed with loop-synchronous polarization scrambling. OF: optical filter, OSA: optical spectrum analyzer, PC: polarization controller ................................................57

Figure 23 Standard deviation of power penalty versus average PDL after 800-km fiber transmission. Solid line: measurement (each solid symbol obtained from 200 samples). Dashed line: estimated results (open symbols) from the experimental results for PMD only and PDL only cases assuming PMD and PDL are statistically independent, i.e. $\sigma_p^2 = \sigma_{PMD}^2 + \sigma_{PDL}^2$ ..................................59

Figure 24 Cumulative probability distribution of the measured power penalties. (a) PMD = 0 (PDL only), and (b) average PMD = 18 ps ..............................................60

Figure 25 (a) Modeling results for variation of Q-factor penalties at the probability level of 2% (square) or 0.1% (circle) as a function of average link PDL. Open symbol: modeling only the effect of signal distortion; solid symbol: the impact of PDL on ASE noise is also considered. (b) Cumulative probability distribution of Q-factor penalties for a link with 2.5-db average PDL and 18-ps average PMD ..............................................................................................62

Figure 26 Simulated probability distribution of the instantaneous PDL for an 8-section link with distributed, 1-dB average PDL .............................................65

Figure 27 Experimental setup for investigating the system impact of polarization scrambling in the presence of PDL for the case of lumped PDL .................66

Figure 28 Power penalty as a function of received OSNR for a group of PDL values with a scrambling frequency of 20 kHz .............................................67

Figure 29 Recirculating loop testbed that can emulate the distributed PDL along an 8×100-km link .........................................................................................68

Figure 30 (a) Statistics of Q factor fluctuation with different PDL values (solid symbols indicate the case without scrambling while open symbols indicates with scrambling case), each curve obtained from 200 samples. (b) Standard deviation of Q factor as a function of PDL per loop with each point obtained from 200 measurement samples ....................................................70
Figure 31 Experimental setup for measurements at 10, 20, 40 Gb/s data rates. AM: amplitude modulator; PM: phase modulator; EAM: electro-absorption modulator; LC-FBG: linearly chirped fiber Bragg grating..............................72

Figure 32 Standard deviation of power penalty as a function of peak-to-peak ripple for NRZ, 50-ps and 10-ps RZ data at 10-Gb/s data rate. Power penalty for different peak-to-peak ripple of NRZ format inserted.................................75

Figure 33 Power penalty distributions for 10/20/40 Gb/s data rates within the 0.8 nm grating bandwidth (grating has 47-ps peak-to-peak ripple).............................76

Figure 34 Power penalty variation of cascading the same FBG using a recirculating loop testbed.................................................................................77

Figure 35 Cascaded FBGs with different gratings that have peak-to-peak ripple of 103-ps (#1), 80-ps (#2), 72-ps (#3) and 47-ps (#4). (a). Solid line: measured ripple after cascading; dashed line: the sum of ripple peaks (worst case in systems) (b). Cascading penalties .........................................................78

Figure 36 Effects of optical filtering after optimum detuning (α=4): (a) Gaussian filter (open square: 10-GHz bandwidth with 8-GHz detuning; solid dot: 15-GHz bandwidth with 14-GHz detuning; solid triangle: 20-GHz bandwidth with 18-GHz detuning) (b) 3-order super-Gaussian filter (open square: 10-GHz bandwidth with 9-GHz detuning; solid dot: 30-GHz bandwidth with 20-GHz detuning; solid triangle: 50-GHz bandwidth with 30-GHz detuning) .....88

Figure 37 Detuning sensitivity of filtering with different bandwidth after 35-km transmission (a) Gaussian filter (b) super-Gaussian filter .......................................89

Figure 38 Filter bandwidth effects for DMLs with different chirp factor: (a) Performance (Q-factor) as a function of bandwidth for different chirp factors. (b) Best bandwidth of the filter for different chirp values .........................90

Figure 39 Performance variation using all-pass filter: (a) a second-order phase response (typical curve of –300 ps/nm) and (b) Q factor as a function of transmission distance for different all-pass filters with different dispersion values ........................................................................................................91

Figure 40 Performance variation (improvement) when we cascade two all-pass filters (the first one has a second-order phase response with a dispersion of 700-ps/nm and the second one has a fourth order at –3000 ps/nm^3) and shift the center frequency of the second filter.................................................................91

Figure 41 Transmission profile and chromatic dispersion of the optical filter used in the experiment........................................................................................................91

Figure 42 Experimental setup of single link transmission without dispersion compensation: For filtering at the transmitter, OF1 is placed after DML; while for filtering at the receiver, OF1 is placed before the receiver (after the transmission fiber). Inserted eye diagrams compare cases with and without optical filtering ........................................................................................................94

Figure 43 Back-to-back Q-Factor improvement using optical filtering under different bias conditions.................................................................94
Figure 44 (a) Experimental results for optical filtering at the transmitter (after DML), two bias conditions are shown here (-0.9 V and -1.4 V); (b) Comparison of optical filtering at the transmitter and the receiver (bias=-1.4 V) ...96

Figure 45 Comparison of optical spectrum and the corresponding eye diagrams for center filtering (maximum optical power) and side-band (asymmetric) filtering: ~ 20-GHz detuning from the carrier center is applied here ...96

Figure 46 Experimental setup of long-haul DML transmission using a recirculating loop testbed. ...98

Figure 47 (a) Optical spectrum of 8x10-Gb/s WDM system after ~1100-km transmission. (b) Relative location of optical filter (upper spectrum) to the DML carrier center (lower spectrum) ...99

Figure 48 Comparison of optical spectra with inserted eye diagrams after ~940-km transmission (a) the optical spectrum of filtering at the carrier center (closed eye) (b) the optical spectrum of asymmetric filtering (eye remains open) ...100

Figure 49 Experimental results showing system performance improvement using asymmetric narrowband filtering. (a) Q-factor of eight channels after 1100-km transmission (under ~ 0-ps/nm residual dispersion). (b) Overall transmission performance (Q vs. distance) of DML under different residual dispersion values ...101

Figure 50 (a) Experimental setup of 40-Gb/s DML transmission. (b) Chromatic dispersion and transmission spectrum of optical filter with 3-dB bandwidth of 0.65-nm ...102

Figure 51 Bit-error-ratio performance of optical filtering effects (a) Transmission of NDF with and without optical filtering (Note that without optical filtering, error-free cannot be achieved even at back-to-back case). (b) Optical filtering with different bandwidths at back-to-back case ...103

Figure 52 Eye diagrams of DML transmission (a) back-to-back without optical filtering; (b) back-to-back with optical filtering and (c) after 25-km NDF transmission using optical filtering ...104

Figure 53 Optical spectrum of filtering (a) comparison of spectrum with and without filtering; (b) Slightly detuning or optimization is necessary when the link length or dispersion varies (here the length changes from 11-km to 25-km) ...104

Figure 54 Experimental setup for measuring the chirp parameter of optical modulators using partial side-band filtering and phase comparison ...107

Figure 55 Analytical result between phase difference and the chirp parameter ...110

Figure 56 Phase difference angle $\theta$ as a function of modulation depth for different chirp parameter values ...111

Figure 57 Comparison of measurement results (phase difference versus chirp parameter at the working bias voltage, i.e. $\tan\theta=1$) and computer simulation results for a dual-drive electro-optic LiNbO$_3$ modulator. Here an F-P filter with ~ 8 GHz bandwidth and 4.7 GHz detuning is used ...112
Figure 58 The corresponding phase differences and output power at different bias voltages when the driving voltage applied to one arm of a dual-drive modulator .......................................................................................................113

Figure 59 Conceptual diagram of chirp measurement using optical spectrum analysis: optical intensity variation of the first sideband of the signal corresponds to the chirp value. In the case of a dual-drive modulator, the 10-GHz RF signal is applied on both arms with an appropriate time delay ......117

Figure 60 Comparison between chirp measurement results using our method and calculated from $(V_1-V_2)/(V_1+V_2)$ method. Here the chirp value is varied by using a dual-drive modulator and changing the driving voltage on one arm………………………………………………………………………….118

Figure 61 Optical spectrum for different chirp values with (a) $\alpha_c=0$, (b) $\alpha_c=0.33$ and (c) $\alpha_c=1.0$. OSA resolution is set at 0.01 nm........................................119

Figure 62 Effect of the signal amplitude on the measurement results of chirp parameter using a commercial modulator with a typical spectrum inserted as the chirp value $\sim 0.75$...............................................................................................120

Figure 63 Tunable pulse-width generation using chromatic dispersion chirped signal .................................................................................................................................................124

Figure 64 Experimental setup for tunable pulse-width RZ transmission using a recirculating loop testbed. AM: amplitude modulator; PM: phase modulator; GEQ: gain equalizer; OF: optical filter..........................................................126

Figure 65 Dispersion vs. wavelength of the sampled nonlinearly chirped FBG ...126

Figure 66 Optical spectrum after 8-loop transmission...........................................127

Figure 67 (a) Power penalties after 600-km single channel transmission with different pulse widths for different residual dispersion values. (b) Transmission distances at $\sim 5$ dB power penalty by tuning the pulse width for different residual dispersion values..............................................................129

Figure 68 Performance of different pulse widths of a typical channel (1555.0 nm) in a 4x10-Gb/s WDM system with 0.8 nm channel spacing under different link residual dispersion values as: (a) 0.4 ps/nm/km (b) 0.08 ps/nm/km. (c) $-0.2$ ps/nm/km. (solid circle: 50-ps pulse; open circle: 35-ps pulse; solid square: 25-ps pulse) .......................................................................................130

Figure 69 (a) Configuration of 3-section fiber squeezer polarization controller and (b) the corresponding polarization state traces on the Poincare sphere as driving voltages vary..................................................................................................................134

Figure 70 Comparison of the DOP reductions from periodic random and three-function scrambling.................................................................................................................................137

Figure 71 Output SOPs of 1000 samples using (a) periodic uniformly random polarization scrambling and (b) three-function polarization scrambling showing “holes” on the sphere.................................................................139

Figure 72 Uniformity of the Stokes parameters for 1000 output SOP samples by (a) three-function scrambling, (b) pure random scrambling with all three
phase retardations uniformly distributed between \([-\pi, +\pi]\), and (c) our proposed approach .................................................................140

Figure 73 (a) Measured output SOPs of ~1000 samples on the Poincaré sphere. (b) Distribution of the Stokes parameters .................................................................141

Figure 74 System applications of uniformly-distributed polarization scrambling: (a) PDL monitoring and compensation (b) PMD monitoring and compensation .................................................................145

Figure 75 Schematic diagram of PDL in-line monitoring and compensation. PC: polarization controller; OF: optical filter; PD: photodiode .........................................................148

Figure 76 The root-mean-square (RMS) power variation of the monitored signal after passing through a 0.8-dB PDL component as a function of the polarization scrambling frequency .................................................................149

Figure 77 Illustration of multiple-channel compensation as the center channel is completely compensated .................................................................150

Figure 78 (a) System demonstration of multi-channel PDL compensation using a loop testbed with close reproduction of the statistics of polarization effects (For single-channel compensation, the scrambling PC is put after the modulator). (b) Measured instantaneous PDL distribution range as well as the average PDL value as a function of transmission distance (100 samples for each distance). .....................................................................................153

Figure 79 (a) Cumulative probability distributions of the power penalty for three different cases: (i) 14-ps average PMD (5.4 ps/loop) with no PDL, (ii) 2.4-dB average PDL (0.85 dB/loop) with no PMD, and (iii) both PMD and PDL combined as in cases (i) and (ii). (b) Cumulative probability distributions of the power penalty after PDL compensation for two typical channels (1556 nm and 1558 nm). Each case in (a) and (b) obtained from 500 samples .................154

Figure 80 (a) Illustration of a programmable DGD module based on polarization switching. (b) Typical response curve of a magnetooptic (MO) polarization switch .............................................................................................................160

Figure 81 Measured DGD and the second-order PMD as a function of designed DGD ......................................................................................................................163

Figure 82 (a) Insertion loss of the delay module as a function of wavelength (1535~1560 nm) at a typical DGD state (~ 40 ps). (b) Polarization dependent loss (PDL) as a function of designed DGD values .................................164

Figure 83 True time delay measurement setup using an ASE broadband source. OSA-optical spectrum analyzer ..................................................................................165

Figure 84 Comparison between the measured absolute true time delay and designed values ..........................................................................................................166

Figure 85 Repeatable DGD generation is used to generate dynamic first-order PMD distribution with tunable average values. (a) First-order PMD distribution of 500 samples with an average DGD of 10 ps. (b) Instantaneous DGD values of the samples ...........................................................................................................167
Figure 86 Characterization of transient DGD effects during polarization switching.
(a) Quasi-static measurement of transient DGD by step-varying the control current of the MO switches (b) Measurement of transient DGD by equivalent small polarization variation between different DGD values from 0 to 45 ps

Figure 87 System evaluation of the tunable DGD module. (a) Experimental setup of a 10-Gb/s NRZ (2^{31}-1 PRBS) link. (b) Measured power penalties for two cases. Open circle: the DGD module remains static after changing state; solid circle: the DGD module keeps jogging back and forth between neighboring states at 1 kHz

Figure 88 Different approaches for PMD compensation using variable DGD modules

Figure 89 Output channel spacing as a function of DGD for a multi-wavelength fiber ring laser

Figure 90 Illustration of data bit alignment using the delay module

Figure 91 Illustration of using the variable TTD module for steering a phased array radar. The relative delays between the radiation elements are adjusted by the variable TTD to cause the total radiation wave front to tilt and hence effectively scan the angle of the radar beam

Figure 92 Illustration of using the variable TTD module in a transversal filter. The output signal is the coherent summation of all branches. The relative phase (delay time) and amplitude changes in the branches alter the shape of the output signal

Figure 93 PMD emulator with three programmable DGD elements separated by two electrically-driven polarization controllers

Figure 94 Three output DGD distributions with different statistical averages, each showing a good fit to the Maxwellian pdf (probability density function) expected from a real fiber, (a) linear scale and (b) logarithmic scale

Figure 95 The corresponding 2nd-order PMD distributions to Figure 94, (a) linear scale and (b) logarithmic scale. The solid curves are drawn from computer simulation results of this emulator, showing that the mean 2nd-order PMD is about 30% lower than that expected from a real fiber with the same average DGD since only 3 sections are used

Figure 96 (a) Output SOP stability of the emulator over 4 hours. (b) and (c) measurements of the DGD and 2nd-order PMD over 30 minutes for two different average DGD values: (b) \( <DGD> = 71 \text{ ps} \), (c) \( <DGD> = 5 \text{ ps} \)

Figure 97 (a) Repeatability of the output SOP as the emulator is repeatedly tuned to 5 DGD states. (b) Four repeated measurements of 50 DGD samples are overlaid to show the repeatability of the output DGD (variation is typically <5% at each sample point with some of this variation due to measurement error)

Figure 98 Emulator setup incorporating simple, in-line polarimeters with automatic feedback to control the SOP between sections. This enables the
generation of a lookup table of input DGD and SOP values versus the
corresponding output 1st ($\tau$) and 2nd-order PMD ($\tau_0$) ..................190

Figure 99 (a) Using polarimeters and feedback control to dial-in previously
recorded PMD states regardless of environmental perturbations (lookup table
concept). (b) Coupling angle between sections set to 0° to align the PMD
vectors and obtain any desired, pure 1st-order PMD and (c) setting the angle
to 90° between sections to obtain large 2nd-order PMDs......................192

Figure 100 Conceptual diagram of PMD emulation using importance sampling,
which is accomplished by applying a biased DGD distribution to each section
(chosen to emphasize the region of interest) and then appropriately weighting
the results to obtain the desired pdf .........................................................193

Figure 101 Importance sampling results for 1000 uniformly distributed DGDs
applied to each section (0 to 45 ps). (a) Measured output pdf (note several
values at large DGDs are generated), (b) pdf after renormalizing the data to
obtain a Maxwellian distribution with $<\text{DGD}> = 15$ ps. Points down to $10^{-24}$
are generated with only 1000 samples. The inset shows the linear scale ....195

Figure 102 2nd-order PMD pdf for the importance sampling experiment described
in Figure 101 (a) Measured output pdf and (b) pdf after renormalizing the
data. The 3-section emulator produces a 2nd-order pdf with the correct shape,
but a slightly lower average than that of a real fiber because of the small
number of sections. The inset shows the linear scale .........................195

Figure 103 Multiple importance sampling results. (a) The three DGD distributions
applied to each section (840 samples/distribution). (b) Resulting DGD
distribution showing that each pdf generates samples in different regions to
cover the entire Maxwellian. (c) Resulting 2nd-order PMD pdf. The insets
show the pdfs on a linear scale where it is evident that better coverage of
lower values is achieved in comparison to the case shown in Figs. 101 and
102.............................................................................................................197

Figure 104 (a) The DGD distributions applied to each section for the
measurements of Q degradation: unbiased (Maxwellian, with average DGD ~
8.7 ps per section) and biased (uniform). (b) Measured Q probabilities using
multiple IS for a system with average DGD = 15 ps ($\sim 10^{-17}$ prob. that Q drops
to $\sim$12 dB) .............................................................................................199

Figure 105 Diagram of proposed DGD monitor ...........................................204

Figure 106 DGD monitoring results (a) Monitored DGD versus normalized power
fluctuation for 10-Gb/s NRZ and RZ data formats using a variable DGD
element. (b) Comparison between the DGD values of the 12-section PMD
emulator given by our monitor (solid circles) and a commercial PMD
analyzer (dashed line) .............................................................................205

Figure 107 (a) First-order PMD compensation for a 10-Gb/s NRZ channel using
monitored DGD as feed forward information. (b) Compensation results......207

Figure 108 Configuration of OSNR/PMD monitor: one PBS, two photodiodes
(PD$_1$ & PD$_2$) and two electrical operational amplifiers (G$_1$ & G$_2$) ........210
Figure 109 Comparison between a previous OSNR monitor approach and the new approach (a) monitoring results using only one polarizer (i.e. one arm of the PBS), PMD has significant effect on the OSNR monitoring results; and (b) OSNR monitoring results of the proposed approach using polarization beam splitting approach (PMD-independent)..........................................................213

Figure 110 PMD monitoring results: (a) depolarization factor ($\delta$) in eq. 6 as DGD $\sim$ 70 ps and 44 ps under different OSNR values and (b) depolarization factor as a function of link DGD ..............................................................................214

Figure 111 (a)Schematic diagram illustrating the concept of proposed higher-order PMD compensation. (b) The corresponding system configuration of higher-order PMD compensation ..........................................................218

Figure 112 Simulation results for 40-Gb/s NRZ transmission:(a) Probability of Q-penalties exceeding 2.5 dB. For the higher-order compensation, the open circles are calculated from the compensation based on feedback control and the solid circles are based on global optimization. (b) Outage probability as a function of Q penalty for first-order compensation and higher-order compensation with 12.5-ps average link PMD ..............................................221

Figure 113 10-Gb/s NRZ experimental setup for higher-order PMD compensation using emulator that consists of 30-section PM fiber with an average PMD of 50 ps. AM: amplitude modulator; PM: phase modulator; PC (1,2): polarization controller; VDGD: variable differential group delay ................222

Figure 114 First-order and higher-order PMD compensation results as a function of (a) DGD value (first-order PMD) (b) second-order PMD.........................223

Figure 115 Adaptive eye mapping (a) Operating principle (MPU: microprocessor unit), (b) typical eye diagram and (c) its corresponding BER contour plot at different sampling points across the data eye ................................................228

Figure 116 Performance evaluation in terms of OSNR (a) comparison of power penalties as a function of OSNR. (b) BER variation across the eye diagram along the vertical direction (see Figure 115(b)) as OSNR changes for the adaptive receiver ..........................................................230

Figure 117 (a) configuration of loop setup; (b) transmission BER curves after ~ 830 km SMF transmission with eye diagram inserted (open symbol: back-to-back; solid symbol: after transmission); (c) performance comparison between conventional receiver and our adaptive transceiver, distance has been doubled from ~ 1000 km to ~ 2000 km ........................................................................................................231
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>autocorrelation function</td>
</tr>
<tr>
<td>AM</td>
<td>amplitude modulation</td>
</tr>
<tr>
<td>BER</td>
<td>bit-error rate</td>
</tr>
<tr>
<td>BW</td>
<td>bandwidth</td>
</tr>
<tr>
<td>CDL</td>
<td>control dependent loss</td>
</tr>
<tr>
<td>CDR</td>
<td>clock data recovery</td>
</tr>
<tr>
<td>CSRZ</td>
<td>carrier suppressed return to zero</td>
</tr>
<tr>
<td>DCF</td>
<td>dispersion-compensating fiber</td>
</tr>
<tr>
<td>DFE</td>
<td>decision feedback equalizers</td>
</tr>
<tr>
<td>DGD</td>
<td>differential-group-delay</td>
</tr>
<tr>
<td>DM</td>
<td>directly modulated (or direct modulation)</td>
</tr>
<tr>
<td>DML</td>
<td>directly modulated laser</td>
</tr>
<tr>
<td>DOP</td>
<td>degree of polarization</td>
</tr>
<tr>
<td>DPSK</td>
<td>differential phase-shift-keyed</td>
</tr>
<tr>
<td>DRA</td>
<td>distributed raman amplifier</td>
</tr>
<tr>
<td>DSF</td>
<td>dispersion shifted fibers</td>
</tr>
<tr>
<td>DUT</td>
<td>device under test</td>
</tr>
<tr>
<td>EA</td>
<td>electro-absorption</td>
</tr>
<tr>
<td>ECSF</td>
<td>equalized carrier-sideband filtering</td>
</tr>
<tr>
<td>EDFA</td>
<td>erbium-doped fiber amplifier</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Term</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>EO</td>
<td>electro-optic</td>
</tr>
<tr>
<td>FBG</td>
<td>fiber Bragg grating</td>
</tr>
<tr>
<td>FEC</td>
<td>forward-error-correction</td>
</tr>
<tr>
<td>FPF</td>
<td>Fabry-Perot filter</td>
</tr>
<tr>
<td>FWHM</td>
<td>full-width half-maximum</td>
</tr>
<tr>
<td>FWM</td>
<td>four wave mixing</td>
</tr>
<tr>
<td>HNL</td>
<td>highly nonlinear</td>
</tr>
<tr>
<td>IS</td>
<td>importance sampling</td>
</tr>
<tr>
<td>ISI</td>
<td>inter-symbol-interference</td>
</tr>
<tr>
<td>JME</td>
<td>Jones matrix eigenanalysis</td>
</tr>
<tr>
<td>LAN</td>
<td>local area network</td>
</tr>
<tr>
<td>LCFBG</td>
<td>linearly chirped fiber Bragg grating</td>
</tr>
<tr>
<td>LCP</td>
<td>left circularly polarized</td>
</tr>
<tr>
<td>LED</td>
<td>light-emitting diode</td>
</tr>
<tr>
<td>LHP</td>
<td>linear horizontally polarized</td>
</tr>
<tr>
<td>LVP</td>
<td>linear vertically polarized</td>
</tr>
<tr>
<td>MLD</td>
<td>maximum likelihood sequence detection</td>
</tr>
<tr>
<td>MZI</td>
<td>Mach-Zehnder interferometer</td>
</tr>
<tr>
<td>NRZ</td>
<td>non-return-to-zero</td>
</tr>
<tr>
<td>NZDSF</td>
<td>non-zero dispersion shifted fiber</td>
</tr>
<tr>
<td>OOK</td>
<td>on-off-keying</td>
</tr>
<tr>
<td>OSA</td>
<td>optical spectrum analyzer</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>OSNR</td>
<td>optical-signal-to-noise-ratio</td>
</tr>
<tr>
<td>PA</td>
<td>Poincaré Analysis</td>
</tr>
<tr>
<td>PBS</td>
<td>polarization beam splitter</td>
</tr>
<tr>
<td>PC</td>
<td>polarization controller</td>
</tr>
<tr>
<td>PCD</td>
<td>polarization-dependent chromatic dispersion</td>
</tr>
<tr>
<td>PDF</td>
<td>probability density function</td>
</tr>
<tr>
<td>PDG</td>
<td>polarization dependent gain</td>
</tr>
<tr>
<td>PDL</td>
<td>polarization dependent loss</td>
</tr>
<tr>
<td>PER</td>
<td>pseudo error rate</td>
</tr>
<tr>
<td>PHB</td>
<td>polarization hole-burning</td>
</tr>
<tr>
<td>PM</td>
<td>phase modulation</td>
</tr>
<tr>
<td>PM(F)</td>
<td>polarization-maintaining (fiber)</td>
</tr>
<tr>
<td>PMD</td>
<td>polarization mode dispersion</td>
</tr>
<tr>
<td>PMDC</td>
<td>PMD compensators</td>
</tr>
<tr>
<td>PS</td>
<td>Poincaré sphere</td>
</tr>
<tr>
<td>PSP</td>
<td>principal state of polarization</td>
</tr>
<tr>
<td>RCP</td>
<td>right circularly polarized</td>
</tr>
<tr>
<td>RZ</td>
<td>return to zero</td>
</tr>
<tr>
<td>SBS</td>
<td>stimulated brillouin scattering</td>
</tr>
<tr>
<td>SCM</td>
<td>subcarrier multiplexing</td>
</tr>
<tr>
<td>SMF</td>
<td>single mode fiber</td>
</tr>
<tr>
<td>SNR</td>
<td>signal-to-noise ratio</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>SOP</td>
<td>state of polarization</td>
</tr>
<tr>
<td>SPM</td>
<td>self phase modulation</td>
</tr>
<tr>
<td>SRS</td>
<td>stimulated Raman scattering</td>
</tr>
<tr>
<td>SSB</td>
<td>single-sideband</td>
</tr>
<tr>
<td>TF</td>
<td>transversal filter</td>
</tr>
<tr>
<td>TTD</td>
<td>true-time-delay</td>
</tr>
<tr>
<td>WDM</td>
<td>wavelength-division multiplexing</td>
</tr>
<tr>
<td>XPM</td>
<td>cross phase modulation</td>
</tr>
</tbody>
</table>
Abstract

Optical signals experience various degrading effects in optical fiber transmission systems and reconfigurable networks. Among those the most important effects include chromatic dispersion, fiber nonlinearities, polarization-related impairments such as polarization mode dispersion (PMD) and polarization dependent loss (PDL), as well as their mutual or combined effects. Depending upon the signal nature (analog or digital, as well as different data formats) and system’s structure (reconfigurable network or point-to-point transmission), the way to improve or optimize system performance may be different.

Recirculating fiber loop testbed is a powerful tool to investigate different perspectives in typical optically amplified long-haul or ultra-long-haul (ULH) systems by sending optical signals through a certain length fiber repeatedly. Most of the study in this dissertation is based on the testbed with different system parameters or configurations.

Polarization-related impairments became a big hurdle to high performance systems, especially when the data rate increases to >10-Gb/s/channel. Both PMD and PDL can induce statistical system fluctuations, and enhanced degradations will happen due to their interaction. Using a loop-synchronized polarization scrambling technique, we replicate a distributed long-haul link with accurate...
polarization statistics. Thus we are able to show experimentally two major degrading effects due to PMD and PDL: (i) combined effects of PMD and PDL, and (ii) degrading effects due to low frequency polarization scrambling in the presence of PDL.

The major part of this dissertation is performance optimization where we show different approaches to (i) overcome the limitations due to fiber dispersion and nonlinearities using pulse-width management for return-to-zero (RZ) systems, (ii) combat PMD and PDL effects by applying dynamic monitoring, optical compensation and electronic mitigation. The advanced methods of PMD emulation, including both the variable first-order emulator and all-order emulator with tunable statistics, have been presented with a further emphasis of experimental realization of importance Sampling.

Although external modulation is the dominating technique in long-haul systems, we also investigate the feasibilities of direct modulation into 10-Gb/s long-haul transmission systems and 40-Gb/s systems using asymmetric narrow-band optical filtering at the transmitter or the receiver.
1 Introduction

Optical fiber communications has been growing rapidly over the past few decades. In addition to the invention of low-loss transmission fiber at two transparent windows (1310-nm and 1550-nm), the major break-through in optical fiber transmission came after invention of Erbium-Doped Fiber Amplifier (EDFA). Due to the wide gain bandwidth of the EDFA, the wavelength-division multiplexing (WDM) channels can be simultaneously amplified and transmitted over long distances. The bit rates have reached 1.28-Tb/s over 70 km for single-channel [1], and 3-Tb/s (300 × 11.6-Gb/s, C+L band) over 7,380 km, 1.28-Tb/s (32 × 40-Gb/s, C band) over 4500 km, and 10.2-Tb/s (256 × 42.7-Gb/s, C+L band) over 100 km for WDM systems [2,3,4]. The spectral efficiency has reached 1.6 (bits/s)/Hz [5].

The demand for network bandwidth is outpacing even the astounding advances of recent years. The ever-increasing fiber optic base and the acceptance of WDM as an established technology are waiting to fulfill the enormous future potential of next-generation networks. Optical networks that go to entirely new lengths are being developed. These networks are being equipped with ultra-long-haul (ULH) technology and fiber-optic lines. Because of its high capacity and performance, optical fiber communications have already replaced many conventional communication systems in point-to-point transmission and networks and also have been considered as a good candidate for wireless backbone.
Natural first step towards optical transport networking is optical channel (wavelength) level reconfiguration, grooming and rapid protection/restoration. The goal is to provide a flexible, scalable, and robust optical transport network, catering to an expanding variety of client signals with equally varied service requirements. The challenges include managing optical channels, optical layer OAM (Operations, Administration, Maintenance), optical layer protection and automated provisioning & distributed restoration.

Current research has explained how to mitigate bandwidth-limiting effects for network capacity up to or more than 1-Tbit/s. However, in the not-too-far future speeds approaching 10-Tbit/s and eventually 10-Tbit/s will be required as WDM optical networks are implemented widely. For those capacities, all the degrading effects on the optical signals in the fiber or devices need to be addressed.

To fully utilize the bandwidth of the fiber and achieve a high performance, several problems need to be solved. Optical signals suffer from many degrading effects in the fiber; including chromatic dispersion, fiber nonlinearities, polarization mode dispersion (PMD), and polarization dependent loss (PDL). Although, these effects and possible solutions have been investigated both theoretically and experimentally for a while, still a lot of issues need to be addressed and better solutions need to implemented to achieve ultra-high capacity and performance systems. Once there is a solution, there are also some challenges emerged from the solution. Thus system
designers are moving forward in a zigzag way by finding the solutions and conquering the challenges.

Currently, scientists and researchers all over the world are trying to combat problems that are more devastating to their applications. In order to mitigate or compensate these degrading effects, appropriate monitoring schemes are necessary as facilitators. On the other hand, efficient emulation tools in both theory and experiments are crucial for performance evaluation and improvement. In this dissertation, I will show not only the new degrading effects, but also and more importantly, how we can further improve the performance and push back the limits for most of the mentioned problems by incorporating emulation, monitoring, mitigation and compensation.
2 Background

2.1 Fiber Dispersion

When an electromagnetic wave propagates in a dielectric waveguide (i.e., silica fiber), the medium response, in general, depends upon optical frequency $\omega$. This property, referred to as chromatic dispersion, manifests through the frequency dependence of the refractive index $n(\omega)$ \[6,7,8,9\]. Fiber dispersion plays a critical role in propagation of optical pulses since different spectral components associated with the pulse travel at different speeds given by $c/n(\omega)$. Consequently, the optical pulse at the output of the fiber will be distorted. Dispersion effect can be considered by expanding the mode-propagation constant $\beta$ in a Taylor series around the center frequency $\omega_0$ \[6\]

$$\beta(\omega) = n(\omega) \frac{\omega}{c} = \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2} \beta_2 (\omega - \omega_0)^2 + .....$$

Where

$$\beta_m = \left[ \frac{d^m \beta}{d \omega^m} \right]_{\omega = \omega_0} \quad (m = 0,1,2,.....) \quad (2.1)$$

The pulse envelope travels at the group velocity ($v_g = 1/\beta_1$), while the parameter $\beta_2$ is responsible for pulse broadening to the first order. $\beta_2$ also depends on the wavelength (i.e., frequency) of the optical signal. The wavelength for which $\beta_2 = 0$ is often referred to as dispersion-zero wavelength ($\lambda_0$); $\lambda_0 = 1.3$ $\mu$m for standard single mode fiber (SMF). More commonly used system parameter is the dispersion
parameter D, and is stated in ps/(nm-km); D for SMF is \(\sim+17.5\) ps/(nm-km) at 1.5 \(\mu\)m. The quantity D is related to \(\beta_2\) by the relation

\[
D = \frac{d\beta}{d\lambda} = -\frac{2\pi c}{\lambda^2} \beta_2
\]  

(2.2)

Dispersion parameters can be changed by tailoring the waveguide profile. In Dispersion Shifted Fibers (DSF), \(\lambda_0\) is in the neighborhood of 1.5 \(\mu\)m and D usually between -2.5 and +2.5 ps/nm-km at 1.5 \(\mu\)m. The dispersion parameter, D as a function of wavelength for both SMF and DSF is shown in Figure 1. Negative D values are referred as normal dispersion (\(\beta_2\) is positive), and positive D values are referred anomalous dispersion (\(\beta_2\) is negative). The wavelength dependency of D is usually considered through dispersion slope that is \(dD/d\lambda \cong 0.08\) ps/nm\(^2\)-km (for both SMF and DSF around 1.5 \(\mu\)m).

![Figure 1 Dispersion parameter D versus wavelength of SMF and DSF](image-url)
Both negative and positive dispersion cause pulse broadening at the output of the fiber. The broadening increases with the fiber length, imposing a limit on the maximum distance and/or data rate without regeneration. Therefore, chromatic dispersion must be mitigated for high-speed or long-distance systems. However, it is apparent by now that chromatic dispersion must be managed, rather than eliminated. Even though it is possible to manufacture fiber with zero dispersion, it is not practical to use such fiber for WDM transmission, because of large four wave mixing (FWM) induced penalties. In theory, compensation of chromatic dispersion for high-speed or long-distance systems can be fixed in value, mostly using dispersion-compensating-fiber (DCF). However, there are several important aspects of optical systems and networks that make tunable dispersion compensation solutions attractive, especially in high-speed optical networks.

In addition, in the WDM + EDFA systems, the dispersion of optical fiber can vary significantly over the EDFA gain bandwidth. In conventional fibers, the slope (dD/dλ) of D is 0.08 ps/nm²/km. This wavelength dependence of chromatic dispersion is labeled second-order dispersion (i.e. dispersion slope) and is important in long-haul WDM systems because different wavelengths may need different dispersion compensation.
Chromatic dispersion arises from the frequency-dependent propagation speed in an optical fiber and is one of the major factors limiting the high-speed performance of fiber-optic communication systems. Although fiber exists zero chromatic dispersion wavelengths, it should be emphasized that such fiber is incompatible with the deployment of WDM systems since harmful nonlinear effects would be generated. As long as WDM is dominant in the marketplace, chromatic dispersion must exist and therefore must be compensated.

Because of the non-zero spectral width of modulated data (optical pulse stream), dispersion leads to pulse broadening, proportional to the distance and with the data rate, thus imposing a limit on the maximum distance transmission without regeneration (see Figure 2a). Dispersion-limited distance can be approximated by determining the transmission distance at which a pulse is broadened by one bit interval. The estimated dispersion limited distance \( L_D \) for a signal having non-return to zero (NRZ) intensity modulation can be obtained by

\[
D \frac{R \lambda^2}{c} L_D = T = \frac{1}{R} \Rightarrow L_D = \frac{c}{\lambda^2 DR^2}
\]

where \( R \) is the data rate, \( T (= 1/R) \) is the bit time, \( c \) is speed of light, and \( \lambda \) is the wavelength of the optical signal. The dispersion limited distance decreases as square of bit rate. For example, the dispersion tolerance of 40-Gb/s systems is
reduced to only 1/16 to that of 10-Gb/s systems. One criterion for detecting this limit for an externally modulated NRZ signal is:

\[ B^2 DL \leq 104,000 (Gb/s)^2 \cdot ps/nm \]  

which corresponds to a dispersion-induced power penalty of 1dB. For single-mode fibers with \( D = 17.5 \text{ ps/nm/km} \), the maximum distance is approximately 1000 km for a bit rate of \( B = 2.5-\text{Gb/s} \), but decreases to about 60 km for \( B = 10-\text{Gb/s} \) and 5 km for \( B = 40-\text{Gb/s} \) in an externally modulated system. Some method of dispersion compensation must be employed for a system to operate beyond these distance limits. In addition, chromatic dispersion of optical fiber is temperature dependent. As shown in Figure 2(b), although the wavelength of zero-dispersion changes slightly with temperature (~ 0.03 nm/°C), it still can become a significant effect for long distance and high data rate systems (> 40-Gb/s) when there is a dispersion slope.
Figure 2 (a) Chromatic dispersion induces optical pulse broadening, proportional to the transmission distance and with the data rate (bits/s). (b) Temperature sensitivity of chromatic dispersion.

The optical eye pattern of a chromatic dispersion limited signal exhibits the effects by closure of the eye. Figure 3 shows a sample eye pattern at different chromatic dispersion value for 10-Gb/s systems. As the fiber transmission distance becomes longer, chromatic dispersion becomes higher, broadening of the pulse becomes greater, causing the eye to close. The power penalty for uncompensated dispersion increases rapidly (exponentially) with the distance.

Figure 3 Sample eye pattern at different chromatic dispersion value for 10-Gb/s system
For \( \geq 10 \)-Gb/s data rates that are transmitted over \( \geq 100 \) km, it is essential that chromatic dispersion be well managed by using some type of dispersion compensation. In theory, compensation of chromatic dispersion for high-speed or long-distance systems can be fixed in value. However, static, fixed dispersion compensation is inadequate when system conditions can change in the following scenarios: (i) reconfigurable optical networks for which a given channel's accumulated dispersion will change when the network routing path is reconfigured, and (ii) \( \geq 40 \)-Gb/s long-distance links for which chromatic dispersion and signal degradation may change substantially due to normal changes in temperature [6]. The required accuracy in dispersion compensation increases dramatically with the data rate. While the amount of residual dispersion that is tolerable at 10-Gb/s is large, of the order of 1000 ps/nm, in 40-Gb/s systems this margin shrinks to only 60 ps/nm. Thus the use of tunable modules is the only way of managing accumulated dispersion.

### 2.2 Fiber Nonlinearities

Most of the nonlinear effects in optical fiber come from the nonlinearity of the index of refraction. This phenomenon is a result of intensity (optical power) dependency of the refractive index. Mathematically, the refractive index of the fiber is related to the optical power as [9]
\[ \tilde{n}(\omega, P) = n(\omega) + n_2 \frac{P}{A_{eff}} \]  

where \( n(\omega) \) is the linear part of refractive index, \( P \) is the optical power inside the fiber, and \( n_2 \) is the nonlinear-index coefficient for silica fibers. The typical value of \( n_2 \) is \( 2.6 \times 10^{-20} \text{ m}^2/\text{W} \). This number takes into account the averaging of the polarization states of the light as it travels in the fiber. The intensity dependence of the refractive index gives rise to three major effects [6, 12]: (i) self-phase modulation (SPM), (ii) cross-phase modulation (XPM), and (iii) four wave mixing (FWM). All these three nonlinear effects can significantly degrade the performance of a WDM lightwave system [13, 14]. XPM and FWM are more severe in multi-channel WDM systems, while SPM can occur in both single channel and WDM systems. I explain SPM and XPM in more detail since it is more relevant to my research and skip FWM and other nonlinear effect.

### 2.2.1 Self-Phase Modulation

Self-phase modulation (SPM) is the phenomenon where any modulation on the signal power gives rise to modulation of the signal phase and spectral broadening. The nonlinear contribution of the index of refraction due to optical power \( P \) results in a phase change \( \Phi_{NL} \) which for light propagating in a fiber given by [9]

\[ \Phi_{NL} = \gamma P L_{eff} \]  

where the quantities \( \gamma \) and \( L_{eff} \) are defined as
\[ \gamma = \frac{2m_i}{\lambda A_{\text{eff}}} \quad \text{and} \quad L_{\text{eff}} = \frac{1 - e^{-\alpha L}}{\alpha} \]

in which \( A_{\text{eff}} \) is the effective mode area of the fiber, and \( \alpha \) is the fiber attenuation loss. \( L_{\text{eff}} \) is the effective nonlinear length of the fiber that accounts for the fiber loss, and \( \gamma \) is the nonlinear coefficient measured in rad/(km-W). The typical range of values of \( \gamma \) is 10-30 rad/(km-W). Although nonlinear coefficient is small, the lengths and powers that have been made possible by the use of the optical amplifiers (EDFAs) can cause the nonlinear phase large enough to play a significant role in the state-of-the-art lightwave systems.

When an intensity-modulated signal travels through an optical fiber, the peak of the pulse accumulates phase more quickly than the wings due to nonlinear refractive index. This results in a nonlinear chirping of the signal. The SPM induced chirp may interact with dispersion induced chirp and can cause a totally different behavior depending upon positive or negative dispersion values [6].

In the normal dispersion regime (\( D < 0 \)), the SPM induced nonlinear-chirp will add to the dispersion induced linear-chirp, thereby causing not only the enhanced pulse broadening but also distorting the shape of the pulse. In the anomalous dispersion regime (\( D > 0 \)), the SPM induced nonlinear-chirp will tend to partially negate the dispersion induced linear-chirp, thereby slightly reducing the pulse broadening, but
still will distort the pulse shape. Therefore, SPM induced chirp can impose a limitation on bit rate and transmission distance in lightwave systems. The SPM induced chirp is dependent upon the power and the shape of the optical pulse. Therefore, if the power and the shape of the pulse is right, the SPM induced chirp and the dispersion induced chirp can completely negate each other in anomalous dispersion regime (D > 0) [7]. The pulse with the right shape and power is called soliton.

### 2.2.2 Cross-Phase Modulation

Cross-phase modulation (XPM) is the phenomenon in which intensity fluctuations in one channel propagating in the fiber modulate the phase of all the other channels or alternatively all the WDM channels (at different wavelengths) in the fiber modulate the phase of any one channel [6, 15].

In a two-channel system, the frequency chirp in channel #1 due to power fluctuation of both the channels is given by

\[
\Delta B = \frac{d\Phi_{NL}}{dt} = \gamma L_{eff} \frac{dP_1}{dt} + 2\gamma L_{eff} \frac{dP_2}{dt} \tag{2.7}
\]

where \(dP_1/dt\) and \(dP_2/dt\) are time derivative of pulse powers of channel #1 and #2, respectively. The first term on right hand side of Eq. (2.6) is due to SPM, and the second term is due to XPM. Note that the XPM induced chirp is twice as much as that of the SPM induced chirp. Therefore, it appears that XPM can impose more...
severe limitation than SPM for WDM systems because effect is twice as large for each interfering channel, and there can be a lot of interfering channels.

Fiber dispersion plays a significant role in the system impact of XPM. Due to dispersion, pulses at different wavelengths travel with different speeds inside the fiber because of group velocity mismatch. In normal dispersion regime (D < 0), a longer wavelength travels faster while the opposite occurs in the anomalous-dispersion regime (D > 0). This feature leads to a walk-off effect that tends to reduce XPM effect. More specifically, the nonlinear interaction (XPM) between the two pulses ceases to occur when the faster moving pulse has completely walked through the slower moving pulse.

The separation in time between the two pulses is governed by the parameter \( d_{12} \) defined by

\[
d_{12} = \beta_1(\lambda_1) - \beta_1(\lambda_2) = v_g^{-1}(\lambda_1) - v_g^{-1}(\lambda_2)
\]  

(2.8)

where \( \lambda_1 \) and \( \lambda_2 \) are the center wavelengths of the two pulses. For pulses of width \( T_0 \), a walk-off length \( L_W \) can be defined as

\[
L_W = \frac{T_0}{d_{12}}
\]

(2.9)

Significance of \( L_W \) comes from the fact that two pulses at different wavelengths in fiber are able to induce phase modulation only for the length \( L_W \). Furthermore, the maximum interaction between the two pulses will occur only when they exactly
overlap each other. Therefore, XPM effect between the two pulses starts increasing when leading edges of the two pulses coincide, and reaches at its peak when the two pulses exactly overlap each other and then starts decreasing, and finally ceases to occur when the two pulses completely separate from each other.

Walk-off length increases with larger wavelength separation between the two channels. Higher walk-off length between the two pulses can significantly reduce the XPM between them. However, in WDM systems, a number of channels at different wavelengths are transmitted together, with a typical wavelength separation range between 0.8-1.6 nm (very large wavelength separation is not desired because of limited bandwidth of EDFAs). Therefore, there can always be some neighboring channels which can distort the channel performance by inducing XPM. Note that, XPM is reduced in regular SMF (D = +17 ps/nm-km) as compared to DSF (|D| < 2.5 ps/nm-km).

2.2.3 Four-wave mixing (FWM)

Like SPM and XPM, four-wave mixing (FWM) is also generated by the intensity-dependence of refractive index of silica. However, impact on performance of WDM system is completely different. In FWM, the beating between two channels of a WDM system at their difference frequency, modulates the phase of one of the channels at that frequency, generating new tones as sidebands [16]. When three
waves of frequencies $f_i$, $f_j$, and $f_k$ interact through fiber nonlinearity, they generate a wave of frequency

$$f_{ijk} = f_i + f_j - f_k$$  \hspace{1cm} (2.10)

Therefore, three waves give rise to nine new optical waves by FWM. For a WDM system with $N$ channels, the number of FWM products generated is

$$M = \frac{1}{2}(N^3 - N^2)$$  \hspace{1cm} (2.11)

In WDM system with equally spaced channels, most of the product terms generated by FWM fall at the channel frequencies, giving rise to crosstalk. The center channels are more vulnerable to this cross talk since the number of FWM products, which fall on center channels, is higher than those, which fall on end channels \[8\][16]. The efficiency of FWM depends on the channel spacing and the fiber dispersion. Increasing channel spacing or fiber dispersion will reduce mixing efficiency.

High-speed WDM systems require simultaneously high launched power and low dispersion values. This greatly enhances the efficiency of FWM, making FWM the dominant nonlinear effect in WDM lightwave systems. FWM can impose severe limitation on bit rate/channel, transmission distance, and number of WDM channels \[16\].

Dispersion limits the maximum transmission distance and the bit rate. But, the effects of XPM and FWM are reduced by dispersion because dispersion destroys the phase matching conditions. In order to achieve good system performance, it is
important to consider the chromatic dispersion and the nonlinear effects of the transmission fiber together. Dispersion management is a solution for this dilemma: use two different types of fibers having opposite dispersions periodically. The total accumulated dispersion is zero after some distance, but the absolute dispersion is non-zero at all points along the link. The result of this dispersion management scheme is that the total effect of dispersion is negligible for all channels, and non-zero dispersion causes phase mismatch between channels thereby destroying efficient nonlinear interactions.

### 2.2.4 Stimulated Scattering

The nonlinear effects described above are governed by the power dependence of refractive index, and are elastic in the sense that no energy is exchanged between the electromagnetic field and the dielectric medium. A second class of nonlinear effects results from stimulated inelastic scattering in which the optical field transfers part of its energy to the nonlinear medium. Two important nonlinear effects fall in this category [11]: (i) stimulated Raman scattering (SRS), and (ii) stimulated Brillouin scattering (SBS). The main difference between the two is that optical phonons participate in SRS, while acoustic phonons participate in SBS. In a simple quantum-mechanical picture applicable to both SRS and SBS, a photon of the incident field is annihilated to create a photon at a downshifted frequency. The new photon is propagated along the original signal in the same direction in SRS,
while the newly generated photon propagates in the backward direction in SBS. Furthermore, the downshifted frequency range where new photons can be generated is \( \sim 30 \) THz in SRS and only \( \sim 30 \) MHz in SBS. Therefore, SBS does not impose any significant limitations in high-speed (Gb/s systems) digital lightwave systems. However, SRS can impose some limitations on WDM systems because the effect of SRS is to deplete the energy of some channels (higher frequency channels) on behalf of the other channels (low frequency channels). The effect of SRS is not very significant unless the number of channels are more than 100 [13]. On the other hand, SRS can be used for signal amplification in a fiber (so called Raman amplifier). Raman amplifier is becoming more and more cost effective now and is extensively developed in recent years because of some unique features. Indeed, unlike EDFA, Raman amplification can virtually occur at any wavelength by properly choosing the pump wavelength and a large bandwidth can be achieved by combing several pump wavelengths.

2.3 Polarization Effects

2.3.1 Background of Polarization

Polarization of light is one of the key parameters in optical fiber transmission systems. Here we give a brief review of polarization theory, details can be found in most of literatures about polarization.
Consider a monochromatic plane wave of angular frequency $\omega$ and of wave vector $k$, parallel to the $z$-axis. In infinite media, and in the case of homogeneous plane waves, the light can be represented by two orthogonal electric field vectors as follows,

\[
E_x(z,t) = \hat{t}E_{0x} \cos(kz - \omega t - \phi_x) \\
E_y(z,t) = \hat{j}E_{0y} \cos(kz - \omega t - \phi_y)
\]  
(2.12)

The complex vector $\vec{E}$, fully described all the attributes of the light wave. The color (wavelength) of light is determined by the angular frequency $\omega$ and wave vector $k$. The intensity of light $I$ is associated to the $\vec{E}$ vector as $I \sim E_{0x}^2 + E_{0y}^2$. With the knowledge of only $\omega$, $k$ and $I$, the $\vec{E}$ vector can not be uniquely defined, because the phase factors also needs to be known. In defining the polarization aspects of light it is the difference between the two phase factors that become relevant. The state of polarization (SOP) of light, as a description of the vector nature of light is described by the so-called polarization ellipse, which can be observed by following the extremity of the $\vec{E}$ vector in the wave plane. The equation of this ellipse can be deduced from equation (1):

\[
\left(\frac{E_x}{E_{0x}}\right)^2 + \left(\frac{E_y}{E_{0y}}\right)^2 - \frac{2E_x E_y \cos\phi}{E_{0x} E_{0y}} = \sin^2 \phi
\]  
(2.13)

where $\phi = \phi_y - \phi_x$ is the phase difference between these two orthogonal components.
Besides the representation described in equations (2.12) and (2.13), another popular method of description of the SOP is Jones vector representation, which represents the SOP by a complex two-component column vector, and the optical medium by a 2x2 Jones Matrix [17]. In this representation, polarized light, as an electromagnetic wave, can be represented as an electric field vector, called the Jones vector. Since any multiplication of the Jones vector by any complex constant does not modify the state of polarization, it is often convenient to work with normalized Jones vectors. The Jones vector can only be used to describe completely polarized light and can be represented as follows,

$$ E = \begin{bmatrix} E_x e^{i\Phi_x} \\ E_y e^{i\Phi_y} \end{bmatrix} $$

(2.14)

This reduced notation must not hide the fact that only the real components can be observed.

The transformation of the state of polarization is represented by a 2x2 matrix, the Jones matrix. Any normal optical components (no depolarization effect, no polarization dependent loss) can be represented by a Jones matrix. For example, a linear birefringent plate whose slow axis is oriented at 0° can be represented by the Jones matrix as follows:

$$ M(\Phi,0) = \begin{bmatrix} e^{-i\Phi/2} & 0 \\ 0 & e^{i\Phi/2} \end{bmatrix} $$

(2.15)

where $\Phi$ is the retardation of the waveplate, and it is defined as
\[
\Phi = \frac{2m\ell(\Delta n)}{\lambda} \tag{2.16}
\]

where \( \Delta n = (n_e - n_o) \) is the difference between the extraordinary refractive index \( n_e \) and the ordinary refractive index \( n_o \). The term \( \Delta n \) is also called the birefringence of the material.

The Jones matrix for a linear birefringent plate oriented at angle \( \alpha \) can be represented as

\[
M(\Phi, \alpha) = R(-\alpha)M(\Phi, 0)R(\alpha) \tag{2.17}
\]

where \( R(\alpha) \) is the Jones matrix for a rotator:

\[
R(\alpha) = \begin{bmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{bmatrix} \tag{2.18}
\]

Jones matrix representation is particularly interesting because the state of polarization transmitted by a series of optical components can be determined by performing the matrix product in the elementary matrices of each element. Generally, the sequence of association of the elements is known, i.e., \( \{M_1, M_2, \ldots, M_k\} \). It is then sufficient to make the matrix product of the elementary Jones matrices, all expressed in the same Cartesian reference system and in the opposite order with regard to the order the light encounters, i.e.:

\[
M = M_kM_{k-1}\ldots M_2M_1 \tag{2.19}
\]
The elementary elements mentioned above are optical components that possess only one birefringence property, such as linear birefringence or circular birefringence.

Stokes parameters [18], an alternate representation of the states of polarization of light, are all real numbers and can be directly measured. Stokes parameters consists of four real numbers, directly related to the intensity of light:

\[
\begin{align*}
S_0 &= E_x E_x^* + E_y E_y^* = I_x + I_y = I_0 \\
S_1 &= E_x E_y^* - E_y E_x^* = I_x - I_y \\
S_2 &= E_x E_y^* + E_y E_x^* = I_{45^\circ} - I_{-45^\circ} \\
S_3 &= i(E_x E_y^* - E_y E_x^*) = I_{RCP} - I_{LCP}
\end{align*}
\]  (2.20)

As shown above, the Stokes parameters \(S_0, S_1, S_2, \) and \(S_3\) are measurable quantities. \(S_0\) denotes total intensity, \(S_1\) refers to the difference in intensities between horizontal and vertical linearly polarized components, \(S_2\) refers to the difference in intensities between linearly polarized components oriented at \(+45^\circ\) and \(-45^\circ\), and \(S_3\) refers to the difference in intensities between left and right circularly polarized components. Thus, the measure of these intensities allows the unambiguously determination of the states of polarization of light. This is significant because it is the basic principle behind many ellipsometers. On the other hand, the physical meaning of stokes parameters is particularly important in the study of partially polarized light, where the notion of the phase difference between components does not have much of a meaning. From the definition, none of the
parameters can be greater than $S_0$, which is usually normalized to 1. For
completely polarized light, we have $S_1^2 + S_2^2 + S_3^2 = 1$, while for entirely
unpolarized light, $S_1^2 + S_2^2 + S_3^2 = 0$. The degree of polarization is therefore
defined as $\gamma = \sqrt{S_1^2 + S_2^2 + S_3^2} / S_0$.

For completely polarized light, the relation $S_1^2 + S_2^2 + S_3^2 = 1$ always holds, so the
parameters ($S_1$, $S_2$ and $S_3$) can be viewed as the coordinates of a point on a unit
sphere. This sphere is the Poincaré sphere, as shown in Figure 4. On the Poincaré
sphere, each point represents a specific polarization state. The linearly polarized
states are represented by the points on the equator. The left and right circularly
polarized states are represented by the south and north poles, respectively. The left
and light elliptically polarized states are represented by points on the lower and
upper sphere, respectively. It is interesting to note that any pair of antipodal points
on the Poincaré sphere corresponds to states with orthogonal polarization.

This representation is particularly useful to describe the polarization transformation
through anisotropic medium. For example, the evolution of the state of polarization
when light passes through a linear birefringent plate can be visualized as a curve on
the Poincaré sphere. As shown in Figure 5, the trajectory of the evolution of the
state of polarization is a curve on the sphere surface. Suppose the wave plate is
oriented at angle $\phi$ relative to the x-axis; the output state can be obtained by a rotation on the Poincaré sphere about axis $\Omega^L$, which lies on the equatorial plane, oriented at an angle $2\phi$ with respect to the $S_i$ axis. The amount of rotation is determined by the retardation of the wave plate.

Figure 4 The representation of polarization states using the Poincaré sphere.

Figure 5 The evolution of the state of polarization as light passes through a linear birefringent plate.
2.3.2 Polarization Mode Dispersion (PMD)

An optical wave of arbitrary polarization can be represented as the superposition of two orthogonally-polarized modes. In an ideal fiber these two modes are indistinguishable, and have the same propagation constants owing to the cylindrical symmetry of the waveguide. However, in real fibers there is some residual anisotropy due to unintentional circular asymmetry as shown in Figure 6. Fiber asymmetry may be inherent in the fiber from the manufacturing process, or it may be a result of mechanical stress on the deployed fiber. The inherent asymmetries of the fiber are fairly constant over time, while the mechanical stress due to movement of the fiber can vary, resulting in a dynamic aspect to PMD. In either case, the loss of circular symmetry gives rise to two distinct orthogonally-polarized modes with different propagation constants.

![Figure 6 Origin of PMD](image)

The difference in propagation constants (differential phase velocity) of these two modes is responsible for polarization mode dispersion (PMD) in the fiber, and can
be related to the difference in refractive indices between the two orthogonal polarization axes as

$$\beta_o - \beta_e = \omega n_o/c - \omega n_e/c = \omega \Delta n_{\text{eff}}/c$$  \hspace{1cm} (2.21)$$

where $n_0$ and $n_e$ are the effective refractive indices of two orthogonal axes, and $\Delta n_{\text{eff}}$ is the differential index of refraction. The differential index of refraction is a measure of birefringence in the fiber, and is usually between $10^{-7}$ and $10^{-5}$. The differential phase velocity indicated in Eq. 2.21 is accompanied by a difference in the group velocities for the two polarization modes. This differential group delay (DGD) that can limit the bandwidth of the fiber by broadening pulses, leads to PMD. PMD is usually expressed in units of ps/km for a short length of birefringent fiber. Typical PMD values for these lengths of fiber are 0.1 to 1.5 ps/km. This linear length dependence of PMD applies to short fibers (1 m to 1 km) where birefringence is considered to be uniform. However, PMD does not accumulate linearly along a long link of fiber. Instead, because of random variations in the perturbations along a fiber span, PMD in one section may either add to or subtract from another section of the fiber. As a result, PMD in long fiber spans accumulates in a random-walk-like process that leads to a square root of transmission-length dependence [19]. Therefore, PMD is expressed in ps/km$^{1/2}$ in long fiber spans, and the typical PMD parameter ($D_p$) is 0.1 to 10 ps/km$^{1/2}$. The probability of the DGD of a fiber section being a certain value at any particular time follows a Maxwellian distribution (see Figure 7). The probability of $DGD = \Delta \tau$ is given by:
\[
prob(\Delta \tau) = \sqrt{\frac{2}{\pi}} \frac{\Delta \tau^2}{\alpha^3} \exp\left[-\frac{\Delta \tau^2}{2\alpha^2}\right]
\]  

(2.22)

with mean value \(<\Delta \tau> = \sqrt{8/\pi}\alpha\). PMD is usually expressed in ps/km\(^{1/2}\) in long fiber spans, and the typical PMD parameter (D\(\rho\)) is 0.1 to 10 ps/km\(^{1/2}\) [19,20,21]. Additionally, in a cascaded fiber link, there may be many discrete components (i.e., isolators, couplers, wavelength multiplexers) that are polarization dependent due to molecular asymmetry (anisotropy) of the waveguide material. Although PMD caused by polarization dependence of a single component may be negligible, cascaded components may add significant PMD in a long link. The combined PMD-induced broadening in a long link may be up to a few tens of ps, which can degrade systems operating at \(\geq 10\) Gbit/s. In systems operating at \(\geq 40\) Gbit/s, PMD has been proved to be deleterious. In order to enable ultra-fast TDM/WDM communications over long distances of optical fiber, the remaining critical issue of PMD must be addressed and understood.

Figure 7 Distribution of first-order PMD (DGD)
In addition to the time variance of DGD, PMD also varies over wavelengths, known as higher-order PMD. This variance results in an optical dispersion that is a function of both the channel bandwidth and the value of DGD over that bandwidth [19]. Figure 8 is a graphical representation of the effect of PMD (both first- and higher-order) on an optical pulse. The optical pulse and its constituent photons travel from the source, or transmitter, at distance = 0, along the single-mode optical fiber. At some distance after PMD has affected the pulse, the polarized energy is separated by some time (i.e., DGD). If DGD is severe, the receiver at some distance \( L \) cannot accurately decode the optical pulse, and bit errors can result.

Figure 8 Graphical representation of the effect of PMD on an optical pulse

The optical eye pattern of a PMD–limited signal exhibits the effects of DGD by closure of the eye. The effect of the eye closure is caused by the separation of the polarized axes of photons, as the DGD becomes higher, separation becomes greater, and optical pulses start to interfere with each other, causing the eye to close (see Figure 9). If the bit errors caused by PMD are too numerous, then the transmitted information is too corrupt to recover and the transmission link should be considered out of service.
The quantity of bit errors encountered at the receiver is directly influenced by the amount of PMD in a fiber optic transmission span. DGD of this magnitude, in a 10-Gb/s transmission system, can be expected to result in a bit-error rate that is severe enough to cause service problems. Some general rules on limitations of distances caused by PMD are given in table 1.

PMD induced problems can be reduced simply by regeneration, i.e., shortening the optical transmission distance. However, from a network point of view, most long-haul transmission systems are now multi-wavelength or dense wavelength division
multiplexing (DWDM) systems, the regeneration is inefficient and costly. In this application, the transmission link must first be demultiplexed, then regenerated, then multiplexed again. This is a very costly operation compared to the preferred alternative of a multi-wavelength amplifier. From a network and cost perspective, a more efficient method of addressing the PMD problem is to fix the effects of PMD while the transmission is in an optical state, before a receiver tries to decode the bits. A PMD compensator (PMDC), deployed at the destination of the transmission system, can reduce the effects of the dispersion in the fiber and ensure that the optical bits are correctly decoded by the receiver before they are to be routed and switched. Electronic solutions are competitive in terms of cost, while adaptive optical PMDCs are more reliable and efficient to realign and correct the pulses of dispersed optical bits.

2.3.3 Polarization Dependent Loss (PDL)

Polarization Dependent Loss (PDL) is an important parameter for evaluating the extent of received-signal degradation in optical transmission systems [22, 23, 24, 25]. It is the difference between the minimum and the maximum transmission of an optical component versus all possible input polarization states. Many optical components, either passive or active, coupler, filters, switches, attenuators and isolators, as well as optical amplifiers, exhibit PDL. Since monitoring/measurement
and compensation of PDL is one topic in this thesis, which will be discussed in
details later, here I will only give a brief introduction about the PDL measurement.

The polarization dependence of the transmission properties of optical components
has many sources, such as dichroism, fiber bending, angled optical interfaces, and
oblique reflection, etc.

Furthermore, the polarization state is not maintained along a fiber. The evolution of
polarization along a fiber is of a completely statistical nature and, in consequence,
is totally unpredictable. Even if the PDL axis of every component is aligned, this
does not correspond to the minimum or maximum effect on polarization sensitive
transmission. Since PDL effects build up in an uncontrolled manner, PDL can lead
to a degradation of the transmission quality of the fiber-optic link, or even to a
failure of the optical system.

In most situations, PDL in a system should be as small as possible. For example, in
long haul telecommunication system, the existence of PDL causes the power
budget management difficult and deteriorates the system performance. In other
situations, PDL is actually desirable, for example, in case of polarizer and
polarization beam splitter (PBS). These components are designed to achieve the
maximum PDL or polarization extinction ratio.
The PDL is defined according to its definition as follows,

\[
PDL = 10 \cdot \log_{10} \left( \frac{P_{\text{max}}}{P_{\text{min}}} \right)
\]  

Therefore, the measurement of PDL is straightforward. In general, there are two kinds of methods of PDL measurement. In the first method, all possible polarization states are generated with the polarization controller and used to observe the changes of loss at the output power of the Device Under Test (DUT), so that the maximum and minimum loss can be determined. This method requires a polarization scrambler to scramble the incident states of polarization. It is usually referred as the Polarization Scrambler Method. A fully functional polarization scramble is desired for such applications. We will discuss our method of polarization scrambling and its application into in-line PDL monitoring later.

Although the Polarization Scrambler Method is simple and easy to be implemented, the use of this method requires a polarization controller that exhibits very little polarization related power variation, including the PDL and Control Dependent Loss (CDL) or Activation Loss. Because of the increasing demands to control the PDL of the devices to be less than 0.005dB, it is becoming more and more difficult to measure such a small PDL using the polarization scrambling method.

The second kind of method is to apply light of known states of polarization to the DUT, the Muller and/or the Jones matrix analysis is used to calculate the PDL.
Depends on which matrix is used in the calculation, it can be termed as the Muller Matrix Method or the Jones Matrix Method.

There are several advantages associated with this fixed states method. Not only is this method fast since as few as three states are sufficient to obtain the PDL as opposed the very large number of states that are necessary for the scrambling method, but the measurements can also be made at many different wavelengths rather rapidly. By the use of this fixed states method, another advantage is that it does not require the ultra small polarization dependent power variation. Because the polarization dependent power variation of a polarization controller is usually caused by systematic changes of the optical system, these repeatable “errors” can be easily removed by subtracting the background reference.

The Muller Matrix Method requires four known incident states of polarization while the Jones Matrix Method only requires three. However, the Jones Matrix Method requires a polarimeter while the Muller Matrix Method requires only a photo detector. Since the polarization scrambling method is quite straightforward, we will only describe the other two methods here.
2.3.3.1 Muller Matrix Method

The Muller Matrix Method determines PDL by exposing the DUT to only four, but well-defined states of polarization (SOP). There are two possible variations of this technique.

The first method, the fixed polarization method, requires a polarization synthesizer. Normally the four SOPs used are: (1, 1, 0, 0), (1, -1, 0, 0), (1, 0, 0, 1). By measuring the intensity response of these incident polarizations, the first row of the Muller Matrix of the system can be determined, and the PDL follows.

Suppose the Muller matrix of DUT has a form

\[
M_{DUT} = \begin{bmatrix}
m_{00} & m_{01} & m_{02} & m_{03} \\
m_{10} & m_{11} & m_{12} & m_{13} \\
m_{20} & m_{21} & m_{22} & m_{23} \\
m_{30} & m_{31} & m_{32} & m_{33}
\end{bmatrix}
\] (2.24)

The polarization transformation of the input state of polarization (SOP) can be calculated by multiplying this Muller matrix \( M_{DUT} \) with the incident Stokes vector \( \mathbf{S} = (S_{i0}, S_{i1}, S_{i2}, S_{i3})^T \). The output power can be represented as follows,

\[
P_{out} = S_{o0} = m_{00}S_{i0} + m_{01}S_{i1} + m_{02}S_{i2} + m_{03}S_{i3}
\] (2.25)

For totally polarized light, we have a constrain that

\[
S_{i0}^2 = S_{i1}^2 + S_{i2}^2 + S_{i3}^2
\] (2.26)

Under this constrain, the maximum and minimum transmission can be expressed as
Therefore, the PDL of the DUT can be easily determined according to equation (2.23).

The second method, the non-fixed polarization method, uses the same principle. Instead of generate the fixed states of polarization using a polarization synthesizer, four states of polarization that are not linear dependent and are reasonably away from each other on the Poincaré Sphere, can be used provided it is repeatable between the reference and measurement. Furthermore, this method requires the use of a polarimeter.

### 2.3.3.2 Jones Matrix Method

Although based on the similar principle, contrary to the power detection in the Muller Matrix Method, Jones Matrix Method relies on the precise measurement of the state of polarization. Once the Jones Matrix is obtained, the PDL is ready to be determined from the eigenvalues and eigenvectors of the product of the Jones Matrix and its complex conjugate and transpose matrix.

As described in the previous sections, for an ideal optical system (i.e. no polarization dependent intensity fluctuation), the Jones Matrix is a unitary complex matrix. The complex conjugate and transpose matrix is equal to its inverse matrix.
Therefore, for an ideal optical system, the product between the Jones Matrix and its complex conjugate and transpose matrix yield a unit matrix. However, if the system exhibits polarization effects such as PDL, this product will not yield a unit matrix; instead, it depends on the amplitude and direction of the PDL vector.

The measured Jones Matrix is a 2x2 complex matrix. The Jones Matrix and its complex conjugate and transpose matrix can be represented as follows,

$$\mathbf{J} = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}, \quad \mathbf{J}^* = \begin{pmatrix} J_{11}^* & J_{12}^* \\ J_{21}^* & J_{22}^* \end{pmatrix}$$

The products of these two matrices is

$$\mathbf{M} = \mathbf{J}^* \cdot \mathbf{J} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

The solutions of characteristic equation yield the eigenvalues of the product matrix:

$$\gamma_{1,2} = \frac{m_{11} + m_{22}}{2} \pm \sqrt{\left(\frac{m_{11} + m_{22}}{2}\right)^2 - m_{11}m_{22} + m_{12}m_{21}}$$

The PDL is calculated from the ratio between these to eigenvalues:

$$PDL = 20 \cdot \log \left| \frac{\gamma_1}{\gamma_2} \right|$$

The coefficient of 20 comes from the fact that Jones matrix deals with the amplitude of the electric field. The power of the light field is the square of the amplitude. The eigenvectors of matrix M are the two states of polarization that exhibit the maximum and minimum loss.
2.4 Recirculating Loop Testbed

Optical loop experiments were performed as early as 1977 to study pulse propagation in multi-mode fiber [24], jitter accumulation in digital fiber systems [25], optical soliton pulse propagation [26], and pulse propagation in single-mode fiber [27]. Loop transmission experiments became useful for optical amplifier feasibility demonstrations after techniques were developed to measure the error performance of long pseudo-random data patterns [28]. The potential transmission capacity of long systems that use EDFA repeaters were understood early on [29][30]; however, experimental verification would take several years from the first amplifier demonstrations. Most of the early loop experiments were performed to simulate the transmission in undersea systems, later diversified to long-haul or ultra-long-haul terrestrial systems. For undersea cable applications, the untested optical amplifier technology was at odds with the strict design requirements of submarine systems. Developing an EDFA cable system would involve extensive resources, both people and equipments. Before resources could be committed, an experimental proof that could demonstrate high data rate, long-length transmission at low BER was needed.

Today, loop techniques play an importance role in the development of long-haul transmission systems by providing a flexible platform for transmission measurements [31], including BER, Q-factor, eye patterns, spectrum analysis, and
optical signal-to-noise-ratio (OSNR), etc. Since the time of the initial feasibility studies using circulating loop experiments, several long testbed experiments have been reported, where hundreds of amplifiers were concatenated to form amplifier chains up to 10,000 km long as mentioned in the introduction part.

Figure 10 Block diagram of recirculating loop transmission, showing the transmitter/receiver pair, transmitter/loop switches, amplifier chain and the error detection circuitry.

Figure 11 Optical switch states for (a) load state and (b) loop state
Typical loop transmission experiment (Figure 10) contains an optical transmitter/receiver pair, a chain of optical amplifier, and performance evaluation equipment (bit-error-ratio test equipment, BERT). Most importantly, optical switching is added in the loop configuration to allow data to flow into the loop (the load state) or to allow the data to circulate (the loop state, see Figure 11).

At a certain time (determined by the round trip time of the closed loop), the control circuitry set the transmitter switch on (or transmitting light) and the loop switch off (or blocking light). The two switches are held in the load condition for at least one loop time to fill the loop with the optical data signal. Once the loop is loaded with data, the switches change state to the loop configuration, and the data is allowed to circulate around the loop for some specified number of revolutions. A portion of the data signal is coupled to the receiver for analysis. The data signal is received and re-timed by the receiver (including O/E converter, electrical amplifier and clock-data-recovery unit, etc) and compared to the transmitted signal in the BERT for error detection. The error signal from the BERT is combined with the error gate in a logic AND gate so that only the errors in the last circulation are counted. The measurement continues, switching between the load and the loop states so that errors can be accumulated over long intervals of time. The BER is calculated as the number of errors detected in the error gate period divided by the total number of bits transmitted during the observation period.
Above description is cited from [31], detailed explanation can be found in the reference. Here I would only emphasize this powerful tool to investigate the system performance for long-haul systems or networks, which is one of the major parts of my Ph.D. thesis.

2.5 Summary

In this section, some of the most important degrading effect in optical signal transmission systems are introduced, including CD, nonlinearities, PMD, PDL, etc., as well as the characterization/measurement of these effects. In addition, as a powerful tool for simulating and investigation of long-haul transmission systems, the configuration and operation principle of the recirculating fiber loop testbed is also described.
3 Investigation of degrading effects

Optical fiber transmission has been extensively used in both analog and digital signals transmissions. Both point-to-point signal transmission and reconfigurable networks have already taken advantage of optics using different signal types (digital/analog) and data formats. However, as described in the previous section, there are many degrading effects in different link configurations. More critically, some mutual or combining effects have been noticed and started from theoretical and experimental investigations. In following, I will show some recent investigation results for various degrading effects.

3.1 Loop-Synchronous Polarization Scrambling Technique for Simulating Polarization Effects using Recirculating Fiber Loops

As described in section 2.4, for the last decade, recirculating fiber loops have been powerful tools in the development of medium-to-long-haul optical transmission systems [32]. A fiber loop experiment can simulate optical signal transmission over thousands-of-kilometers by recirculating the signal through an optical amplifier chain of modest length (i.e., hundreds of kilometers), and therefore provides a highly economic way to investigate various fiber transmission issues [31]. The recirculating-loop technique has been well developed to accurately replicate the
optical noise, fiber dispersion, and nonlinear effects in a point-to-point link. However, conventional recirculating loops may become inadequate in the presence of non-negligible lightwave polarization effects, specifically, PMD and PDL.

As we know, the first-order PMD, i.e. DGD, is a random variable that has a Maxwellian probability density function, and the mean DGD increases with the square root of distance. A critical aspect of investigating PMD induced penalties is to account for the low probability but high degradation tail of this distribution. Because recirculating loops have been widely used as experimental testbeds for installed systems, it is highly desirable that loop testbeds can correctly reproduce the statistical distribution of polarization effects in straight-line systems. However, the periodic nature of fiber loop may artificially produces an unrealistic PMD distribution that is skewed towards higher DGD values [33]. For the same reason, the statistical distribution of PDL effect in a recirculating loop may be significantly different from that in a straight-line system [34].

In this section, we will describe the demonstration of a loop-synchronous polarization scrambling technique that can make the distributions of polarization effects in a recirculating loop similar to those in a straight-line system [33], [35].
3.1.1 Loop-Synchronous Polarization Scrambling

Figure 12 shows the recirculating fiber loop for experimental demonstration. The non-return-to-zero (NRZ) 10-Gb/s transmitter consists of an external-cavity laser followed by a Mach-Zehnder modulator. The dispersion-managed fiber loop consists of three EDFAs operating in the saturated regime, 82 km of conventional SMF, and 12 km of DCF. The SMF has a chromatic dispersion of 17 ps/km/nm and a loss of 0.25 dB/km. The residual chromatic dispersion of the loop is +45 ps/nm at the signal wavelength (1557nm). The input powers to the SMF and DCF are 3.0 dBm and -1.0 dBm, respectively. PMD can be introduced by including a short piece of polarization-maintaining (PM) fiber of a fixed DGD value in the loop. Without this PM-fiber, the entire loop has a background PMD (average value) of 1.0 ps. Optical switching is used to load data into the loop and then allow data to circulate through the loop.
To emulate the random polarization evolution in a straight-line fiber link, a loop-synchronous polarization controller (LSPC) is inserted into the loop. The transmission matrix (Jones matrix) of the LSPC is determined by the loop control circuitry and updated after each round trip of the optical signal circulating through the loop. This technique, referred to as loop-synchronous polarization scrambling, is illustrated in Figure 13. In loop experiments, the basic time unit is the round-trip time of the loop $\tau_{\text{loop}}$, which is 0.46 ms here. The entire loop running cycle contains multiple round-trip intervals for data-loading and recirculation. In Figure 13, the loading time is $2 \times \tau_{\text{loop}}$, but it can be reduced to $1 \times \tau_{\text{loop}}$. Without using any technique of polarization scrambling within the loop, the same polarization transformation is applied to the signal after each round trip through the loop; hence, we can manually adjust the polarization controller so that the closed loop has a unit transmission matrix and the loop output SOP maintains constant (see trace (a) in Figure 13). The periodic polarization transformation is eliminated when the LSPC generates a series of random, uncorrelated Jones matrixes during each load-and-circulating cycle. Two types of commercial polarization controllers have been tried in our experiment. One is a LiNbO$_3$ device having a fast transition time of $< 1$ µs, the other is based on piezoelectric fiber squeezers with a transition time of $\sim 50$ µs. For a loop with a single LSPC, the transition timing of the LSPC matches the error bursts that occur on the seams of the loops [31].
Figure 13. Oscilloscope traces illustrating the state-of-polarization (SOP) of the loop output signal. (a) Without any polarization scrambling inside the loop, a constant output SOP can be obtained by manually adjusting a polarization controller (PC) inside the loop. (b) The SOP is randomized using the loop-synchronous polarization controller, which repeatedly generates a series of random transmission matrixes during every loading-and-circulating cycle. We used a simplified polarimeter consisting of a polarizer followed by a O/E converter of 100 kHz bandwidth to measure an arbitrary component of the SOP vector. The oscilloscope displays the output photocurrent. The gating signal for burst-mode BER measurement is also shown.

The polarization transfer condition of the circulating loop is shown in Figure 14, where $P_{k,j}(\omega)$ is the transmission matrix of the LSPC, $N$ is the number of circulations, $T_{k}(\omega)$ is the frequency-dependent transmission matrix of the remainder part of the loop. Here, the $k$-subscript indicates a statistical sample of the fiber link to be emulated. In ordinary lab environment, the link PMD (or PDL) condition may maintain tens of seconds if the LSPC repeatedly generates the same series of Jones matrices in each load-and-circulating cycle. Therefore, a bit-error-
rate (BER) as low as $10^{-9}$ can be measured for $\geq 10$-Gb/s data. Good repeatability of the polarization controller is required for low BER measurement, since the data has to be measured and accumulated over many loop cycles. The random polarization transformation $P_{ij}$ is expressed as

$$P = \begin{bmatrix}
\sqrt{\gamma} \exp(i\theta) & -\sqrt{1-\gamma} \exp(-i\phi) \\
\sqrt{1-\gamma} \exp(i\phi) & \sqrt{\gamma} \exp(-i\theta)
\end{bmatrix}$$

(3.1)

where $0 \leq \gamma \leq 1$, $-\pi \leq \theta \leq \pi$, and $-\pi \leq \phi \leq \pi$ are statistically independent and uniformly-distributed random variables. The control voltages of the LSPC are fully determined by the $(\gamma, \theta, \phi)$ parameters produced from a random number generator.

Note that, given an arbitrary $2 \times 2$ unitary matrix $U$, the matrix product $UP$ or $PU$ has identical distribution as $P$ (see Appendix for a proof). Thus, the random transformation $P$ scatters the SOP uniformly on the Poincaré sphere for any input SOP.

![Diagram](image)

**Figure 14.** Polarization transfer condition of the recirculating loop shown in Figure 12.

Polarization scrambling within a fiber loop has been reported in recent loop transmission experiments, where the polarization scrambler runs independently on the loop control circuitry [36][37]. Generally, such asynchronous polarization-
scrambling scheme results in a measurement of BER that is averaged over many realizations of the link illustrated in Figure 14. However, the distribution of system performance may be measured if the system operates under a very poor ($>2 \times 10^{-4}$) BER [37]. In this case, the BER for $>10$ Gb/s data can be determined within a $<5$ µs interval that may be much shorter than the time scale on which the system polarization condition changes.

### 3.1.2 Experimental Results of PMD Emulation

Figure 15 shows the histogram of received optical power at $10^{-9}$ BER for the system in Figure 12, where we used a LiNbO$_3$ polarization controller and a short piece of PM-fiber of 8.3-ps DGD. The loop PDL is $\sim 0.3$ dB and is mainly introduced by the LiNbO$_3$-PC. We use 8 circulations through the loop, corresponding to the transmission over a fiber link of 772 km. The received optical signal-to-noise-ratio (OSNR) in 0.1-nm bandwidth is $\sim 26$ dB. In the back-to-back case, the receiver reaches BER floor at $10^{-9}$ for an OSNR of 21.5 dB. With loop-synchronous polarization scrambling, the transmission matrix for each round-trip of the loop is randomized. As a result, the root-mean-square (RMS) DGD of the link is $\sim 23.5$ ps and the average link DGD is $\sim 22$ ps. We measured the receiver sensitivity for 1000 independent realizations of the link. There are 2% of the samples leading to a power penalty exceeding 3 dB as compared to the case without intentionally introducing PMD into the system. We also made
measurement when the \( \text{LiNbO}_3 \)-PC is turned off and set at a random state. For this case without polarization scrambling inside the loop, the link DGD is skewed towards much higher values than that may occur in a straight-line system. As a result, the distribution tail of power penalties extends to >10 dB, and error floors (>\(10^{-9}\) BER) are observed for 8% of the samples.

The average link DGD value can be easily changed by altering the DGD of the short PM-fiber inside the loop. Figure 16 shows the distribution of power penalties while using a piece of PM fiber of 5.4-ps DGD. In this case, the average link DGD is \(~14\) ps, and the power penalty tail at 2% probability is 1.3 dB. The power penalty at 2% probability is also measured for the case with an average link DGD of 18 ps or 27 ps (Figure 17).
Figure 16. Histogram similar to Figure 15 (a) except that DGD/loop is 5.4 ps.

Figure 17. PMD induced power penalties at $10^{-9}$ BER versus the average DGD values.

To get a reference of PMD-induced penalty for our experimental system, we also measured BER for the case where PMD is introduced outside the loop. The experimental setup is similar to Figure1 except that the PM fiber is moved to
between the receiver-end EDFA and the optical filter. A piece of PM fiber of 43-ps DGD can induce ~ 4.5 dB power penalty in the worst case of 50% polarization splitting ratio (see Figure 18).

![Figure 18. BER curves for (i) back-to-back performance of the PIN receiver used in our experiment, (ii) dispersion managed 772-km transmission with negligible PMD, (iii) transmission over 772-km fiber of negligible PMD followed by a PM fiber of 43-ps DGD (worst-case polarization splitting).](image)

3.1.3 Modeling and Discussion

3.1.3.1 Using single LSPC

We first analyze the statistical distribution of DGD in our recirculating-loop system using the model shown in Figure 14. For the case of loop containing one fixed DGD element, the polarization transmission matrix for the \( j \)-th circulation through the loop (and \( k \)-th sample of the transmission link), denoted by \( L_{j,k}(\omega) \), is given by
where $\Delta \tau_0$ is the DGD per loop, and $U_k$ is a random polarization transform that does not change with loop circulations. Using loop-synchronous polarization scrambling, polarization transforms $P_{i1} \cdots P_{iN}$ are randomized and are independent with each other. The histogram of link DGD for 10000 samples is shown in Figure 19 (a), where we assume the number of circulations $N = 8$. The DGD distribution for this case of $N = 8$ is quite close to Maxwellian. In contrast, Figure 19 (b) shows the distribution of DGD for the case without scrambling the round-trip transmission matrix, so $P_{ik}$ becomes a random transform $P_k$ that will not change with loop circulations. When $P_k U_k$ is approximately identity matrix, the link DGD will add up linearly with loop circulations. Therefore, a very large DGD value can occur at a high probability and the DGD distribution deviates remarkably from the Maxwellian distribution.

Assume DGD-induced penalty is expressed as $e(dB) = A\gamma(1-\gamma)(\Delta \tau / T_b)^2$ [38], where the splitting ratio $\gamma \in [0,1]$ has a uniform distribution, the DGD $\Delta \tau$ has a distribution as shown in Figure 19 (a), $T_b = 100$ ps is the bit period, and the constant $A = 97$ is selected so that 43-ps DGD introduces a maximum penalty of 4.5 dB. We calculate the penalty at 2% probability to be 3.2 dB. This result agrees well with our experiment.
Figure 19 Histogram of link DGD after circulating 8 times through a loop that contains one fixed DGD element. (a) With loop-synchronous polarization scrambling, (b) without polarization scrambling.

3.1.3.2 Using multiple LSPC’s

In general cases, we may employ multiple LSPC’s in a fiber loop, as illustrated in Figure 20. The transition timing of one LSPC corresponds to error burst while the others are delayed by a fraction (i.e. 1/3) of the loop round trip time one after the other. The loop transmission matrix (j-th circulation, k-th sample) is written as

$$ L_{ij}(\omega) = \prod_{m=1}^{M} T_{km}(\omega) P_{ij,m} $$

(3.3)

where $T_{km}(\omega)$ is transmission matrix of the fiber between two LSPC’s. For example, we consider only PMD effect and model $T_{km}(\omega)$ as an intentionally-introduced, PM-fiber concatenated with a long transmission fiber, and we have

$$ T_{km}(\omega) = U_{km} \begin{bmatrix} e^{i \gamma_{km} \tau/2} & \epsilon^{-\gamma_{km} \tau/2} \end{bmatrix} \epsilon^{-\gamma_{km} \tau/2}, \quad \tau_{km} = (\Delta \tau_{0m}^{2} + \Delta \tau_{m}^{2} - 2 \gamma_{km} \Delta \tau_{0m} \Delta \tau_{m})^{1/2} $$

(3.4)
where $\Delta \tau_{0m}$ is the DGD of the PM-fiber (a fixed parameter), $\Delta \tau_{km}$ is the DGD of the transmission fiber (a random variable of Maxwellian distribution), and $\gamma_{km}$ is the polarization coupling ratio (uniformly-distributed between 0 and 1).

Figure 20. Schematic diagram of a recirculating loop using multiple loop-synchronous polarizations controllers.

In the following, we will provide modeling results for a 450-km recirculating loop with three LSPC’s $(M = 3)$. The length of fiber following each LSPC is 150 km. and the PMD of transmission fiber is 0.05 ps/(km)$^{1/2}$. Figure 21 (a) shows the histogram of link DGD ($10^5$ samples) after circulating 8 times through the loop, assuming three pieces of PM-fiber introduced into the loop have DGD values of 1.5, 2.5 and 3.5 ps, respectively. We obtain an accurate reproduction of the Maxwellian distribution, and the simulated probability of $\Delta \tau > 3\Delta \tau$ is exactly $4 \times 10^{-5}$. Under the same assumption of the loop, Figure 21 (b) shows the wavelength
autocorrelation function of the PMD vector [39]. For this case using only three unequal DGD elements, the average level of residual correlation between two well-spaced wavelengths is ~10%. A normal PMD emulator needs 15 unequal DGD sections to achieve such a low residual correlation [40].

\[ \langle \text{DGD} \rangle = 12.3 \text{ ps} \]

Figure 21. (a) Histogram of link DGD after circulating 8 times the loop shown in Figure 13, with \( \Delta \tau_{i0} = 1.5 \text{ ps} \), \( \Delta \tau_{i2} = 2.5 \text{ ps} \), and \( \Delta \tau_{i3} = 3.5 \text{ ps} \). The inset plots the tail of the cumulative distribution function (CDF) in log scale. The exact Maxwellian distribution is plotted in smooth lines. (b) Autocorrelation function of the PMD vector.

A loop of single LSPC is able to reproduce the Maxwellian distribution if (i) the loop circulation number is large enough (i.e. >15) and (ii) the intentionally-introduced, fixed DGD is much larger than the PMD of transmission fiber. However, with multiple LSPC’s, required number of circulations is significantly reduced, and PMD is distributed more uniformly along the link.

Detailed analysis about the polarization matrixes can be found in [41], as well as some other experimental and theoretical results.
3.2 Combined Effects of PMD and PDL in Long-Haul Systems

High-bit-rate (≥10 Gb/s) fiber-optic communication systems are vulnerable to problems arising from various fiber polarization effects, such as PMD and PDL [42]. It is well known that PDL, as well as polarization dependent gain (PDG) of optical amplifiers, can induce random fluctuations of the system signal-to-noise-ratio along the link, which leads to a significant performance degradation in long-distance systems [43]. Recent publications have also showed theoretical analyses that PDL can enhance the pulse broadening that is caused by PMD [44, 45, 46, 47]. However, experimental results of the combined effects of PMD and PDL have not yet been reported.

In this section, we experimentally demonstrate the probability distribution of power penalties of these two polarization effects using a recirculating fiber-loop testbed for 10-Gb/s NRZ transmissions over an 800-km fiber link. We find that the PDL of in-line components combined with the link PMD will increase the penalty distribution. With an average PMD of 18 ps, the system penalty distribution tail at 2% probability increases from 2.5 to 4.3 dB as the link-average PDL increases.
from 1.0 to 2.1 dB. In the absence of PMD (PDL only), we observe much lower power penalties.

### 3.2.1.1 EXPERIMENTAL SETUP

We use a fiber loop testbed to replicate the long-distance optically-amplified link. As described in section 3.1, conventional recirculating loops are inadequate in the presence of non-negligible polarization-dependent effects, for example, PMD. This problem exists because a recirculating loop exhibits some measure of periodic behavior that artificially produces an unrealistic PMD distribution that is skewed towards higher DGDs. Therefore, we employ the loop-synchronous polarization control technique to randomize the transmission matrix of each round-trip of the loop, and to obtain an effective tool for measuring statistically the effects of PMD and/or PDL [41].

Figure 22 shows the experimental setup. A tunable laser at 1557 nm is modulated at 10 Gbit/s (2^{15}-1 PRBS). The loop configuration is similar to the one described in Figure 12. The average PMD of 82-km SMF and 12-km DCF are 0.8 ps and 0.5 ps, respectively. In order to correctly emulate the statistical distribution of PMD and PDL, we use a loop-synchronous LiNbO$_3$ polarization controller (PC), a piece of PM fiber, and a variable PDL element inside the loop. The Jones matrix of the LiNbO$_3$ PC is determined by the loop control circuitry and updated after each
round-trip interval of the loop in order to generate a series of random, uncorrelated polarization states during the entire loop running period (a predefined number of loops). PMD is introduced by including a piece of PM fiber in the loop. The variable PDL element acts as a wavelength-independent partial polarizer that can be adjusted from 0.15 dB to 0.9 dB. The background PDL of the loop is ~0.25 dB and is mainly due to the LiNbO$_3$ PC. We use 8 circulations through the loop, corresponding to ~650-km SMF transmission or ~800-km fiber transmission. The received optical SNR (OSNR in 0.1 nm bandwidth) is ~26 dB.

![Experimental setup of the recirculating loop testbed with loop-synchronous polarization scrambling.](image)

**Figure 22** Experimental setup of the recirculating loop testbed with loop-synchronous polarization scrambling. OF: optical filter, OSA: optical spectrum analyzer, PC: polarization controller.

During the experiment, the system power penalty is determined by comparing the receiver power sensitivity at 10$^{-9}$ BER with the back-to-back sensitivity measured at 26-dB OSNR. The receiver reaches BER floor at 10$^{-9}$ with an OSNR of 21.5 dB. The intrinsic power penalty without the intra-loop PM fiber and PDL element is
~0.7 dB. We estimate the average link PDL as the square-root of the number of loops multiplied by the PDL-per-loop [44], which is confirmed in the later section about PDL monitoring and compensation. For the intra-loop PMD, two spools of PM fiber with 6.9-ps and 8.4-ps DGD are used, corresponding to an average system PMD (8 loops) of 18 ps and 22 ps, respectively.

3.2.1.2 EXPERIMENTAL RESULTS

Figure 23 shows the standard deviation of the system power penalty after 800-km transmission under different conditions of combined PMD and PDL, with each solid data point in the figure being based on 200 sample measurements. The solid lines represent experimental measurements, and the dashed lines represent estimations under the assumption that PMD and PDL effects are statistically independent, i.e. $\sigma^2 = \sigma_{PMD}^2 + \sigma_{PDL}^2$ with $\sigma_{PMD}$ and $\sigma_{PDL}$ the experimental results for PMD-only and PDL-only cases, respectively. Without the intra-loop PM fiber (i.e., PDL only), the standard deviation of power penalties is caused mainly by PDL-induced fluctuations in the received SNR and is a relatively small effect. When PMD is introduced into the loop, the system penalty standard deviation will increase with an increase in PDL. Under the scenario that the link average PDL is lower than 1.0 dB (0.35 dB/loop), the interaction between PMD and PDL is not obvious. However, when PDL is greater than 2.1 dB (0.75 dB/loop), significant performance degradation is observed. In this regime, the power penalty standard
deviation increases from 0.8 dB to 1.9 dB as the average PDL varies from 0.7 dB to 2.5 dB with 22-ps PMD.

Figure 23 Standard deviation of power penalty versus average PDL after 800-km fiber transmission. Solid line: measurement (each solid symbol obtained from 200 samples). Dashed line: estimated results (open symbols) from the experimental results for PMD only and PDL only cases assuming PMD and PDL are statistically independent, i.e. $\sigma^2 = \sigma_{\text{PMD}}^2 + \sigma_{\text{PDL}}^2$.

To further compare system performance under conditions of small and large PDL, Figure 24 shows the probabilities of power penalties in which each probability curve is derived from 500 experimental samples and the figure-of-merit is taken for 98% of all samples. Without incorporating PMD into the loop, Figure 24 (a) shows that the power penalty is <2.5 dB even when the average PDL is increased to 2.1 dB. However, Figure 24 (b) shows that when PMD is introduced, enhanced degradation does occur. If we compare the power penalty distribution tail at 2%
probability, as the average PDL increases from 1.0 dB to 2.1 dB, the 2% tail only changes from 1.6 dB to 2.4 dB without PMD, but it can increase from 2.5 dB to 4.3 dB with 18-ps PMD. Such an increase in power penalty exceeds the result expected under the assumption that the effects of PMD and PDL are statistically independent, showing that the performance degradation is due to the combined effect of PMD and PDL.

![Graph showing cumulative probability distribution of measured power penalties for different PMD and PDL values](image)

**Figure 24** Cumulative probability distribution of the measured power penalties. (a) PMD = 0 (PDL only), and (b) average PMD = 18 ps.

### 3.2.1.3 Modeling Results

We also investigate the combined effect of PMD and PDL based on a Monte Carlo simulation of signal propagation. The simulated link is similar to the system replicated by our loop testbed, except that residual chromatic dispersion, fiber nonlinearity, and the PDG of optical amplifiers are not considered in the simplified
model. We calculate the penalties of Q-factor induced by PMD and PDL, assuming that optical noise (i.e. ASE) is the dominant noise source. The ASE noise is spectrally resolved, with each frequency component represented by a four-dimensional Stokes vector for partially polarized light [43]. Signal (or ASE) propagation is modeled in the frequency domain, using $2 \times 2$ transmission matrixes (or $4 \times 4$ Muller matrixes) to describe polarization-dependent elements. Note that our simulation considers not only the SNR fluctuation but also the pulse-shape distortion due to PMD and PDL.

Figure 25(a) shows the Q-factor penalty distribution tail versus the average link PDL at a probability level of 2% and 0.1%, with the average link PMD maintained at 18 ps. We perform 20,000 simulation cycles for optical links containing 8 DGD and 8 PDL sections. Note that, at a probability level lower than 2%, the DGD distribution in log-scale of the simulated link deviates gradually from Maxwellian. The solid symbols show the Q-penalties as in the real case, where unpolarized ASE is added at the place of each EDFA. Here, system penalties are induced by pulse broadening as well as the enhanced signal-ASE beat noise [43]; the fluctuation of the total ASE power is negligibly small. In order to isolate the effect of pulse distortion, we also present modeling results (open symbols) with unpolarized ASE added only before the O/E converter. In the presence of PDL, the interference between the two non-orthogonal PSPs can either increase or decrease the eye opening. However, our modeling shows that, for systems with small-to-medium
PMD (PMD × bit-rate < 0.18) and an average link PDL < 2dB, the PDL enhanced pulse distortion has a minor effect on the system penalty at a probability as low as $10^{-4}$, and PDL induces system degradation mostly because of the enhanced signal-ASE beat noise.

Figure 25 (a) Modeling results for variation of Q-factor penalties at the probability level of 2% (square) or 0.1% (circle) as a function of average link PDL. Open symbol: modeling only the effect of signal distortion; solid symbol: the impact of PDL on ASE noise is also considered. (b) Cumulative probability distribution of Q-factor penalties for a link with 2.5-dB average PDL and 18-ps average PMD.

Furthermore, Figure 25(b) illustrates the cumulative probability distribution of the Q-factor penalty induced by 2.5-dB average PDL and/or 18-ps average PMD. Without PMD, the Q-penalty (in dB) has a Gaussian distribution with a standard deviation of 0.4 dB. In this case, we can also simulate power penalties based on the data of our receiver sensitivity varied with OSNR. The resulting standard deviation of power penalty due to 2.5-dB PDL is 0.4 dB, which agrees with the experiment. With PMD included, it is not easy to establish a model of the receiver sensitivity
for a distorted signal and directly compare the simulation with the experiment. However, it seems that the modeling produces a somewhat lower penalty in the presence of both PMD and PDL. This needs further investigation in both modeling and experiments.

### 3.3 Deleterious System Effects due to Low Frequency Polarization Scrambling in the Presence of PDL

High-speed polarization scrambling of the signal at the transmitter has been shown to significantly reduce signal fluctuations due to PDG or to reduce PDL-induced noise [48, 49, 50]. In general, polarization scrambling modulates the signal's state-of-polarization (SOP) at a given frequency. Previous research concerning polarization scrambling of a signal has concentrated on its effects in long cascades of erbium doped fiber amplifiers (EDFAs), typically on the order of hundreds of amplifiers. Results have shown the effectiveness of polarization scrambling for ultra-long undersea transmission systems using either low frequency (~10 kHz) or bit-synchronous polarization scrambling [48, 49, 50]. It has been noted that the PDL-induced intensity modulation is a potential system degrading effect [43]. However, it has also been shown that this effect can be dramatically suppressed by the PDG that exists in a cascade of many (e.g. hundreds) saturated EDFAs using low frequency scrambling. PDG becomes effective in suppressing the PDL effect since the scrambled signal gets partially depolarized every time it passes through a
PDL element - PDG acts as feedback that tries to depolarize the signal and minimize the PDL-induced amplitude modulation [43].

Recently, polarization scrambling has also been shown to be useful for PMD compensation since it can provide a means of monitoring the instantaneous differential group delay (DGD) in the feedback tracking loop [51, 52, 53]. For such monitoring, polarization scrambling moves the input SOP to random points on the Poincarè sphere on a millisecond time scale. The scrambling frequency is defined as the SOP update frequency and is typically tens of kHz.

However, no published work has appeared that describes the impact of this kind of low-frequency polarization scrambling for cases in which the instantaneous PDL may be high but the PDG effect is small enough (e.g. in typically terrestrial systems where the number of EDFAs is on the order of tens or less) to be considered negligible. In such cases, the PDL-induced intensity modulation may become a deleterious effect.

In this section, we experimentally demonstrate that low-frequency polarization scrambling at the transmitter will produce significant system performance degradation for a link with non-negligible values of instantaneous PDL. For example, with 2.6-dB instantaneous PDL, the receiver power penalty is ~3 dB at a 26-dB optical signal-to-noise-ratio (OSNR) for 10-Gb/s NRZ transmission. For a
link with distributed, 1-dB average PDL, the probability of an instantaneous link PDL exceeding 2.6 dB is $\sim 10^{-4}$, as shown in Figure 26.

![Figure 26 Simulated probability distribution of the instantaneous PDL for an 8-section link with distributed, 1-dB average PDL.](image)

In the case of a distributed 8-span link using a recirculating loop testbed, when the PDL value of one span (per loop) is less than 0.3 dB, the deleterious effect is not obvious, however, as the PDL value increases, the effect becomes significant. We emphasize that network designers might determine that polarization scrambling will be necessary for PMD monitoring or PDG reduction, but this would require either the reduction of PDL in in-line components or the use of a simple PDL compensator [54].
3.3.1 Lumped PDL-induced degradation

We first use a variable lumped PDL element to investigate the system impact of PDL-induced intensity modulation due to polarization scrambling.

![Experimental setup for investigating the system impact of polarization scrambling in the presence of PDL for the case of lumped PDL.](image)

Figure 27 Experimental setup for investigating the system impact of polarization scrambling in the presence of PDL for the case of lumped PDL.

The experimental setup for 10-Gb/s NRZ data ($2^{23}$-1 PRBS) transmission is shown in Figure 27; some typical eye diagrams are also included for comparison. We choose a fiber squeezer based polarization scrambler due to its superior intrinsic characteristics, such as low loss, low PDL and negligible additional phase modulation, and put it right after the transmitter. The polarization scrambler can generate fast periodic polarization scrambling over a series of uncorrelated SOPs. The first optical attenuator is used to adjust the OSNR of the received signal. The power penalty is compared with the back-to-back sensitivity measured at $10^{-9}$ BER for different OSNR and different PDL values at a 20-kHz scrambling frequency as shown in Figure 28.
Without the PDL element (measurement background), the additional power penalty due to scrambling is < 0.1 dB. Significant performance degradation occurs when a link has non-negligible values of PDL. For example, a 2.6-dB instantaneous PDL gives ~3 dB power penalty at a 26-dB OSNR. We find that when the scrambling frequency is less than 1 kHz, the scrambling effect is negligible since the EDFA operating in the saturated regime is able to alleviate PDL-induced low speed power fluctuations, and the optical receivers usually cut off all fluctuations at a frequency lower than their cut-off frequencies (kHz range). In contrast, if the scrambling frequency is beyond tens of kHz, which is the required frequency for the applications in PMD compensation, the optical intensity modulation will be detected at the receiver and cause eye closure.
3.3.2 Distributed PDL statistics

Although the use of a lumped PDL element can somewhat illustrate the effects of polarization scrambling, it is more interesting to investigate the effect in a real long distance transmission link with distributed PDL statistics. In order to emulate these statistics, we use a recirculating loop testbed that closely reproduces the PDL statistics in a typical terrestrial long distance system, as shown in Figure 29.

A single channel at 1555 nm is NRZ modulated at 10 Gbit/s ($2^{23}-1$ PRBS). A polarization scrambler with 20-kHz scrambling frequency is placed right after the transmitter. The dispersion-managed loop consists of three EDFAs operating in the saturated regime, 82 km of SMF, and 12 km of DCF. In order to emulate the statistical distribution of PDL, the loop contains a loop-synchronized polarization
controller (PC) and a variable PDL element, as described in the previous sections. The background PDL of our loop is ~ 0.25 dB. The PDL element can be changed from 0.05 dB to 0.9 dB. Here we use 8 passes that emulate an 800-km transmission link with the average PDL growing according to the square root law. The received optical SNR without PDL inside the loop is ~27 dB.

For a given PDL value inside the loop, we measure the BER fluctuation at the receiver as the link instantaneous PDL varies, i.e. different measurement samples according to PDL statistics, and then calculate the Q factor using the relationship between BER and Q factor, i.e. $BER = \frac{1}{2} \text{erfc}(\frac{Q}{\sqrt{2}})$. The two cases - with and without polarization scrambling are compared in Figure 30.

In Figure 30 (a), each curve plots the statistical results obtained from 200 measurement samples. With 0.45 dB and 0.9 dB PDL inside loop, which correspond to ~1.3 dB and 2.5 dB average PDL along the whole link, respectively, the 5% Q-factor tails without scrambling are 5.1 and 4.9 (corresponding BER of $1.3\times10^{-7}$ and $4.0\times10^{-7}$, respectively) which decrease to 4.5 and 3.8 with polarization scrambling (corresponding BER of $4.0\times10^{-6}$ and $5.0\times10^{-5}$, respectively). Furthermore, the standard deviations of Q factor as a function of PDL per loop are also compared in Figure 30 (b). Note that with a per-loop (one span) PDL value of
less than 0.3 dB, the effect is not obvious, but as the PDL increases, serious performance fluctuations (higher standard deviation) occur.

![Graph](image)

Figure 30 (a) Statistics of Q factor fluctuation with different PDL values (solid symbols indicate the case without scrambling while open symbols indicate with scrambling case), each curve obtained from 200 samples. (b) Standard deviation of Q factor as a function of PDL per loop with each point obtained from 200 measurement samples.

### 3.4 System Impacts by the Group-Delay Ripple in Single and Cascaded Chirped Fiber Bragg Gratings

For long distance fiber transmission wavelength-division-multiplexed (WDM) systems, nearly all ≥10-Gb/s channels require the periodic management and compensation of fiber chromatic dispersion. The most common type of deployed chromatic dispersion compensator is dispersion compensating fiber (DCF). Another potential alternative is the use of dispersion-compensating chirped fiber Bragg gratings (FBGs), for which the reflection time delay (ps) is a function of
wavelength (nm) [55, 56]. Such gratings have the potential for the following characteristics: (i) fixed or tunable operation, (ii) single- or multiple-channel functionality, (iii) low nonlinearity, (iv) compact footprint, (v) low loss, and (vi) compensation for dispersion slope of the transmission fiber.

Unfortunately, the fabrication process of FBGs is not perfect and produces random "ripples" in the group delay (GD) performance of the grating. These ripples adversely affect the quality of the compensated signal and limit the utility of FBG-based solutions. The effects of ripple depend on the ripple peak-to-peak and root-mean-square (RMS) values as well as the ripple frequency in relation to the signal bandwidth. Previous theoretical investigations of the system effects of ripple have elucidated several issues [57, 58, 59, 60, 61, 62]. However, although some measurements have been published [58, 59], actual system measurements of the relationship between ripple values and system power penalties are still quite interesting. Moreover, detailed measurements need to be reported for the effects due to cascaded FBG compensators, such as would be encountered when using FBGs as periodic in-line compensators.

In this section, we experimentally show impacts of group-delay ripple on system performance when considering a single FBG and cascaded FBGs. We find that, for single gratings with different GD ripple, shorter pulse-widths (smaller duty cycles)
can reduce the penalty induced by GD ripple at the same data rate, meanwhile, the impact of ripple becomes more and more serious as the data rate increases, e.g. the maximum power penalties are ~ 0.9, 1.7, 2.7 dB after dispersion compensation with the same grating (67-ps peak-to-peak ripple) for 10, 20 and 40-Gb/s, respectively. For cascaded FBGs, cascading the same grating results in similar performance as the ripple adds up linearly. However, as confined in our measurements, cascading several gratings with random ripple does not degrade system performance as much since the ripple growth obeys a square root law as the number of gratings increases.

3.4.1 Experimental Setup

Throughout our measurements, unless specified, we use the experimental setup shown in Figure 31.

![Figure 31 Experimental setup for measurements at 10, 20, 40 Gb/s data rates. AM: amplitude modulator; PM: phase modulator; EAM: electro-absorption modulator; LC-FBG: linearly-chirped fiber Bragg grating.](image-url)
At the transmitter, the light from a tunable laser source is first NRZ modulated using an amplitude modulator at 10 Gb/s, then the RZ signal with 0.5 duty cycle is generated by a second amplitude modulator driven by the clock signal. After that, the signal is further compressed down to ~10 ps through a phase modulator followed by 80-ps/nm of SMF. A two-stage optical multiplexer, including a polarization interleaver, is used to generate the 20 and 40Gb/s data streams. Thus if we bypass some of these components at the transmitter, we can get 10-Gb/s NRZ, RZ (with duty cycle 0.5 and 0.1), 20-Gb/s RZ, 40-Gb/s RZ signals, respectively. At the receiver end, an optical feedback loop is used to recover the 10 GHz clock from the data stream, optical demultiplexing from 20 or 40-Gb/s down to 10-Gb/s is realized through an electro-absorption (EA) modulator electrically driven by the recovered 10 GHz clock (For 10-Gb/s data, the EA modulator is bypassed). An optical pre-amplifier is used before the receiver to increase its sensitivity. The power penalty measured at $10^{-9}$ bit-error-rate (BER) after transmission over SMF and the grating compensator is compared with the back-to-back sensitivity. The fiber length is varied according to the dispersion of the gratings in order to get complete compensation and isolate the effects of ripple itself. The back-to-back sensitivities of NRZ and RZ formats at 10-Gb/s data rate in our system are -32 dBm and -34.5 dBm, respectively. In our experiment, we chose several linearly-chirped FBGs (LC-FBGs) with different GD ripples, but with similar reflectivity amplitude ripple (~1 dB) and GD ripple period (less than 50 pm). Although the period of GD ripple is also an important parameter for assessing performance degradation [57,
it is difficult to compare different periods due to the uncertainty of fabrication. Since the gratings used in our experiments have similar ripple periods, we only consider the magnitude of GD ripple. The PMD values of all these gratings are less than 3 ps over the bandwidth. We measure the penalties over the bandwidth of the gratings by varying the signal wavelength with a less than 0.01 nm step size. The statistical values are then derived from the measured penalties.

In addition, in order to compare with our experimental results, we performed simulations to calculate the power penalties caused by the GD ripple based on the frequency response of the FBG within its passband given by [55]

\[ H(\omega) = e^{j\Phi(\omega)}, \quad \Phi(\omega) = \int \tau_g(\omega)d\omega = -\frac{2\pi c}{\lambda^2} \int \tau_g(\lambda)d\lambda. \quad (3.5) \]

where \( \tau_g(\lambda) \) is the measured group-delay response of the FBG. And our experimental results agree with the performed simulations.

### 3.4.2 Single FBG

First we measure the impact of GD ripple for different pulse widths at a 10-Gb/s data rate, as shown in Figure 32.
Figure 32. Standard deviation of power penalty as a function of peak-to-peak ripple for NRZ, 50-ps and 10-ps RZ data at 10-Gb/s data rate. Power penalty for different peak-to-peak ripple of NRZ format inserted.

Here we compare three cases: NRZ and RZ with 0.5 and 0.1 duty cycles. The result shows that shorter pulse widths lead to smaller performance variations due to the corresponding broader optical spectrum of the signal. In addition, the maximum power penalties of NRZ (similar to RZ) for the gratings with peak-to-peak ripple of 47, 65, 103 and 164 ps are ~ 0.8, 1.0, 1.8 and 5.5 dB, respectively (inserted in Figure 32).

In order to investigate the GD ripple impacts on systems with different data rates, we use a grating with ~ 65-ps peak-to-peak ripple and less than 1-ps PMD across the bandwidth. As shown in Figure 33, within an ~0.8 nm bandwidth, the maximum power penalties for 10, 20 and 40 Gb/s data rates are ~ 0.9, 1.7 and 2.7 dB, respectively, and the corresponding standard deviations are 0.2, 0.36 and 0.51
dB. For the same grating, higher data rates experience not only higher power penalty (although partially due to the residual dispersion), but also larger performance variation over the grating bandwidth.

![Power penalty distributions](image)

**Figure 33** Power penalty distributions for 10/20/40 Gb/s data rates within the 0.8 nm grating bandwidth (grating has 47-ps peak-to-peak ripple).

### 3.4.3 Cascaded FBGs

For long-haul applications, cascaded FBG characterization is important as FBGs can be used as periodic in-line compensators for each span. We first investigate the worst case that cascades the same grating again and again, where the ripple accumulates linearly each time when the signal passes through the grating. Instead of DCF, we put a LC-FBG (47 ps peak-to-peak GD ripple) as the dispersion compensator inside a recirculating fiber loop. The loop (pass) number corresponds to the number of cascading the grating itself.
Figure 34 Power penalty variation of cascading the same FBG using a recirculating loop testbed.

As shown in Figure 34, the maximum power penalty grows quickly as the number of passes increases. After 3-loop transmission where the ripple accumulates to ~150-ps peak-to-peak, the maximum power penalty is ~6 dB, which agrees well with the performance of a single grating with similar ripple (shown in Figure 33).

However, it is more realistic in practical systems that the grating ripple adds randomly since most ripples generated in the fabrication process are random. The peak-to-peak ripples of the four gratings (FBG #1, #2, #3 and #4) used here are 103-ps, 80-ps, 72-ps and 47-ps. The measured ripple after cascading these gratings is shown in Figure 35(a). Both the cascaded peak-to-peak and RMS ripple grow
according to the square root law as the gratings cascade. The corresponding power penalty is shown in Figure 35(b).

![Figure 35](image)

**Figure 35** Cascaded FBGs with different gratings that have peak-to-peak ripple of 103-ps (#1), 80-ps (#2), 72-ps (#3) and 47-ps (#4). (a) Solid line: measured ripple after cascading; dashed line: the sum of ripple peaks (worst case in systems) (b) Cascading penalties.

Compared with the case of Figure 34 where the grating with the smallest ripple is used, the power penalty increases much more slowly when cascading different gratings. The maximum penalty increases from 1.8 dB to only 3.4 dB after cascading all four gratings even though the ripples of the other three gratings are higher than that of the one used in Figure 34. We also note that the above measured power penalties agree well with our simulation results using the measured delay data of the cascaded gratings. Although this effect after cascading gratings with random ripple somewhat relaxes the requirements on grating ripple, there is still some possibility that the ripple of some gratings may be located at the same
wavelength, thus the worst case still needs to be considered. On the other hand, new techniques in fabricating gratings can reduce the ripple significantly, which makes FBGs more promising as in-line dispersion compensators [64].
4 Technical solutions for performance optimization in WDM systems and reconfigurable networks

In section 3, we have investigated several new degrading effects that may bring significant system impacts. On the other hand, optimize system performance by combating degrading effects is more interesting and important for system and network designers. In this section, we will concentrate on different approaches to optimize or facilitate the optimization of system performance.

4.1 Reach Extension in 10-Gb/s and 40-Gb/s Directly Modulated Transmission Systems Using Asymmetric Narrowband Optical Filtering

Direct modulation schemes have attracted increased attention during the past few years due to their intrinsic simplicity and cost-effectiveness, especially when applied to metro and access networks [65], as well as coarse wavelength-division-multiplexing (CWDM) systems [66]. As the two major parameters in the system design, data rate and transmission distance are given attention and efforts from different prospective, maximizing the benefits from conquering the intrinsic chirp of directly modulated lasers (DMLs) and induced spectrum broadening [67]. Especially, it has been assumed that external modulation is required to transmit a 10-Gb/s data signal over long distances. This is unfortunate, given that direct
modulation is exceptionally cost effective compared to external modulation. It has long been a goal in the optical communications community to find a technique that will enable the use of inexpensive direct modulation schemes for long-distance transmission.

At the data rate of 2.5-Gb/s, the intrinsic chirp from DML is not that crucial to the system performance, and the transmission distance can be up to several hundred kilometers [68, 69]. However, when DML is deployed into 10-Gb/s systems, transmission distance is limited to tens of kilometers over single-mode fiber (SMF), which is the typical transmission distance for metro and access networks. In order to enhance the achievable distance, a number of approaches have been used, including (i) mid-span spectral inversion (over 200-km) [70], (ii) electrical equalization in the receiver (20-km) [71], (iii) 0.2-nm filter at high driving current (38.5-km) [72], (iv) the deployment of negative dispersion fiber (100 ~ 320 km) [73, 74]. Dispersion supported transmission (DST) has been used for 20-Gb/s DML with a transmission distance of 53 kilometers [75].

When the data rate increases to 40-Gb/s, a rate that even external modulation has encountered a lot of critical problems, there are only few reported results on 40-Gb/s DML at 1310-nm transmission window for very-short-reach (VSR) applications [76] or extended distance over 40-km using 4-level signal [77]. Very
recently, transmission results at 1550-nm using dispersion-managed links, i.e. SMF with dispersion compensating fiber (DCF), have been reported [78, 79].

Due to the cost and difficulty of (or lack of feasibility in) changing embedded fiber links, a method that enhances system performance while only requiring modifying one or both of the endpoints of a link is a critical requirement. In addition to the approaches mentioned above, modification of the DML incorporating other components or reshaping the corresponding driving current has been demonstrated [78, 80, 81, 82]. Since frequency modulation (FM), as a function of the bit rate, is introduced by the amplitude variation by directly modulation on the DFB laser, using interferometer schemes that can utilize the frequency modulation had also been shown the effectiveness [83, 84], as well as the electrical approach after detection [85].

In this section, we will concentrate on another efficient approach – optical filtering in the 10-Gb/s system, and further apply this technique into 40-Gb/s systems. Optical filtering has been widely used in high performance communication systems to narrow the signal spectrum and increase the spectral efficiency up to 160% [86, 87, 88]. However, in directly modulated systems, depending on the filter characteristics, optical filtering can improve the performance due to two major
mechanisms: narrowing the broadened spectrum [89] and converting FM to useful amplitude modulation (AM) for transmission [90, 91, 92].

First we isolate the contribution of two mechanisms of optical filtering using computer simulation, especially the effect of narrowing spectrum using an optical Gaussian filter. In directly modulated systems, detuned filtering from the carrier center (i.e. asymmetric) is necessary. Narrowing spectrum shows to be more effective in terms of extending the transmission window (i.e. similar performance through a long distance), while FM to AM conversion is simulated using an all-pass filter and shows the similar performance improvement as dispersion compensation or transmission through certain optical fiber [90]. Design guideline is given in terms of filtering bandwidth, detuning and sensitivity, etc.

Based on the simulation conclusion, we investigate the impact of asymmetric optical filtering in a 10-Gb/s directly modulated system using a commercial DML with a wavelength ~1.55-μm. An optical filter with Gaussian amplitude profile and a 3-dB bandwidth of 0.3-nm is used in the experiment. In addition to the narrowing spectrum, chirp management (effective FM to AM conversion) is achieved via the interaction between the nonlinear phase response of the sloping edge of the optical filter, the laser chirp, and the link chromatic dispersion, thereby significantly improving system performance. Experimental results show the effectiveness of optical filtering in doubling the reach from <25 km to >45 km (without dispersion
compensation) even as the bias condition was varied. Similar performance can be obtained by putting the filter at either the transmitter or the receiver side.

Furthermore, we demonstrate error-free transmission (Q > 15.6 dB) at 10-Gb/s up to 1,400 km when using the DML. We use dispersion management along the link incorporating with optical sideband filtering at the receiver to achieve our result. The DML is inserted into an 8x10-Gb/s WDM system. The maximum error-free (Q > 15.6 dB) transmission distances for the DML are ~ 1400 km, 1100 km and 580 km for a link with residual dispersion values of –0.54 ps/nm/km, 0 ps/nm/km and +0.60 ps/nm/km, respectively.

### 4.1.1 Asymmetric narrowband optical filtering

In directly modulated systems, chirp is caused by the change in the refraction index induced by carrier injection during modulation, as well as the power-dependent photon intensity distribution along the laser cavity. Time evolution of wavelength change or frequency chirping may broaden the optical spectrum and degrade signal quality quickly along the fiber link [93]. When an optical filter is applied, two mechanisms will play the corresponding roles:

(i). Compared with external modulation, apparently there are quite amounts of unwanted frequency components in the broadened spectrum due to direct...
modulation. While strong filtering has been used in external modulation systems to increase the spectral efficiency [86, 87, 88], an optical filter with a bandwidth narrower than the broadened spectrum in directly modulated systems may reduce these frequency components, therefore a “cleaner” signal may transmit longer distance. In addition, chirped induced spectrum broadening is not symmetric, consequently the optical filtering will also be asymmetric. Another reason for the asymmetric filtering will be discussed later.

(ii). Fiber grating filter has been shown to improve the frequency response of the DML [90], as well as reduce the intensity noise of the DML [92], later also been applied into chirp reduction of semiconductor optical amplifiers (SOAs) [91]. In these cases, in addition to the pulse reshaping with the help of narrowing spectrum, when the grating filter has pre-designed chirp or dispersion, it will interact with the laser chirp (or even the link dispersion along the fiber link), thus convert the frequency modulation into useful amplitude modulation and add on the transmitted signal, resulting in better performance. Note since most of the chirp or dispersion of optical filter is introduced by the nonlinear phase response on the sloping edge, the filter should also be detuned from the center wavelength of the carrier. For example, a DML with positive chirp may require the center frequency or wavelength to be located at the negative edge of the filter accordingly.
4.1.2 Simulation results

Although the two mechanisms of filtering effects in directly modulated systems have been mentioned separately before [89, 90], investigation of the effects in details, as well as providing useful design guidelines if possible, is still quite beneficial to the community. In order to isolate the two contributions, we use computer simulations that can easily emulate different conditions, such as characteristics of filter and DML, link configurations, etc., specifically, we simulate two kinds of typical optical filters: Gaussian filter and all-pass filter.

In the previous analysis using the trapezoidal filter [94], we found that: (i) The optimized position of the filter depends on the sign and magnitude of the laser chirp and the fiber dispersion map; (ii) The bandwidth of the filter does not have a significant effect on the Q-factor once the ideal detuning frequency and slope shape has been determined. After optimization, it is the edges of filters that matter. Since the trapezoidal filter is not a physically practical one, we choose the Gaussian filter as a candidate that can provide the effect of narrowing spectrum. Ideally the Gaussian filter has linear phase response [95], thus there is no dispersive effect involved. In our simulation, the typical design parameters for the DML is listed in table 2, except that we vary the linewidth enhancement factor $\alpha$. For the Gaussian filter, we change the bandwidth of the filter, and scan the filter (i.e. detune the center frequency of the filter) across the optical spectrum of the signal.
As shown in Figure 36 (a), where we set the enhancement factor $\alpha$ to be 4 (typical values for commercial DMLs are between 2 and 4. Normally as $\alpha$ decreases, the performance increases), and detune the filter with different bandwidths to the optimum position, the transmission distance after filtering can significantly improved compared with the case without filtering (although the best Q-factor is decreased). Within a certain distance (typically ~ 40-50 kilometers), the performance in terms of Q-factor keeps constant, resulting in much wider transmission window for the DML. Since the sloping edge of the Gaussian filter is still not sharp enough, we further apply a super-Gaussian filter with an order of 3 into the simulation for comparison, as shown in Figure 36 (b). Note that super-Gaussian filter has already been used in high capacity transmission systems to
increase the spectral efficiency [88]. The sharper slope induced by super-Gaussian filter may provide a stronger effect for narrowing the spectrum, thus the performance may be slightly better than the normal Gaussian filter.

One of the major concerns for system designers when using narrowband optical filters is the sensitivity of filter position. If the location of the filter is too sensitive, it will be very challenging to stabilize the filter. Figure 37 (a) shows the sensitivity of Gaussian filter as a function of the detuning frequency at a transmission distance of 35 kilometers. Fortunately we can see that there is a relatively flat frequency window with more than 4 GHz even for the narrowest Gaussian filter. Figure 37 (b) shows the case for the super-Gaussian filter. Unlike the Gaussian filter, where the
best performance for filters with different bandwidths after optimum detuning is
similar, the performance of super-Gaussian filters with different bandwidths after
detuning has quite differences due to the sharp sloping edges. In turn, the Gaussian
filter is preferable in terms of stability and sensitivity.

![Graphs showing detuning sensitivity for different bandwidths after 35-km transmission.]

Figure 37 Detuning sensitivity of filtering with different bandwidth after 35-km transmission (a) Gaussian filter (b) super-Gaussian filter

Above results are for the DML with the factor \( \alpha \) of 4. As \( \alpha \) changes, the best choice
for the filter should also change accordingly. For example, Figure 38 (a) shows the
performance comparison of Gaussian filter at different \( \alpha \) values after 35-km
transmission. Subsequently, the best bandwidth for the Gaussian filter at different \( \alpha \)
values is shown in Figure 38 (b). As the chirp factor \( \alpha \) increases, the spectrum
broadening will be more serious, thus the bandwidth of the filter should also
decrease.
Figure 38 Filter bandwidth effects for DMLs with different chirp factor: (a) Performance (Q-factor) as a function of bandwidth for different chirp factors. (b) Best bandwidth of the filter for different chirp values.

Frequency modulation (FM) to amplitude modulation (AM) conversion using filtering has been discussed in [90, 91, 92]. In order to illustrate this effect, we use the all-pass filter as an example. In practical systems, all-pass filter has been used in different applications, especially chromatic dispersion compensation [96]. First we assume the phase response of the filter is second-order (a typical curve shown in Figure 39 (a)), i.e. the induced chromatic dispersion is a constant value. Then we find the performance variation when applying this filter in the DML system, as shown in Figure 39 (b). As expected, this is similar to the effect of a length of transmission fiber with a certain dispersion value. The performance shifts with slightly difference under different dispersion values. However, when we add higher-order phase variation, this linear shift changes. Figure 40 shows the performance variation when we cascade two all-pass filters (the first one has a second-order phase response with a dispersion of 700-ps/nm and the second one
has a fourth order at $-3000$ ps/nm) and shift the center frequency of the second filter. This implies the performance improvement when using all-pass filter in the DM systems, in terms of not only the best transmission distance, but also the wider transmission window (i.e. similar Q-factor at different distances).

![Phase Response and Q-Factor](image)

Figure 39 Performance variation using all-pass filter: (a) a second-order phase response (typical curve of $-300$ ps/nm) and (b) Q factor as a function of transmission distance for different all-pass filters with different dispersion values.

![Performance Variation](image)

Figure 40. Performance variation (improvement) when we cascade two all-pass filters (the first one has a second-order phase response with a dispersion of 700-ps/nm and the second one has a fourth order at $-3000$ ps/nm) and shift the center frequency of the second filter.
In real systems, most of the optical filters (even Gaussian filters) have some amount of chromatic dispersion. This implies that directly modulated systems using optical filtering benefit from the combining effects of both narrowing spectrum and FM to AM conversion.

4.1.3 Single link transmission without dispersion compensation

Under the guidance of simulation results, we choose an optical filter with a Gaussian transmission profile as shown in Figure 41. The filter is a thin film based filter with a 3-dB bandwidth of ~ 0.25 nm (~ 30-GHz). Unlike the Gaussian filter used in the simulation, this filter does have a certain amount of chromatic dispersion cross the passing band (on the magnitude of tens of ps/nm).

Figure 41 Transmission profile and chromatic dispersion of the optical filter used in the experiment.
First we perform single link transmission over SMF without dispersion compensation. The experimental setup is shown in Figure 42. The DML is directly modulated at 9.95328-Gb/s ($2^{23}-1$ PRBS) with an amplitude of 2 V. The laser is driven at 40 mA and we tried a number of bias points. Different lengths of SMF comprise the transmission link. For the case of optical filtering at the transmitter, the optical filter (OF1) is used after the DML; for the case of filtering at the receiver, the filter is placed after the transmission fiber. The receiver consists of an optical attenuator, a pre-amplifier (EDFA2) and another wideband optical filter OF2 (>1-nm). The attenuator is used to keep the input power into the pre-amplifier constant, thus the optical signal-to-noise-ratio (OSNR) will not affect the Q-measurement results. OF2 is used to reduce the ASE noise from optical amplifiers. The input power into the PIN receiver is fixed at ~3 dBm. Q measurements are performed at the receiver using decision threshold adjustment. In the back-to-back condition, we find that through optical filtering, the Q factor can be improved by more than 2 dB at different bias voltages, as shown in Figure 43. In addition, the typical eye diagrams are inserted in the setup for comparison.
Figure 42 Experimental setup of single link transmission without dispersion compensation: For filtering at the transmitter, OF1 is placed after DML, while for filtering at the receiver, OF1 is placed before the receiver (after the transmission fiber). Inserted eye diagrams compare cases with and without optical filtering.

Figure 43. Back-to-back Q-Factor improvement using optical filtering under different bias conditions.

By changing the link length, we compare the performance improvement using optical filtering. Figure 44 (a) shows the experimental results in terms of the transmission distance versus Q factor using optical filtering at the transmitter (after the DML). Two bias conditions are compared here: -0.9 V and -1.4 V. The
maximum transmission distances (with a Q-factor of ~ 15.6 dB used as the threshold for maximum transmission distance) are almost doubled using filtering under these two bias voltages (i.e. from ~ 20-km to >40-km with -0.9-V bias and from ~25-km to >45-km with -1.4-V bias). In addition, we can see that without filtering, although the Q factor at -1.4-V bias is lower than the one at -0.9-V for the back-to-back condition, the -1.4-V bias case results in a longer transmission distance than the one at -0.9-V: the –0.9-V bias case suffers a much more rapid Q-factor degradation as the transmission distances increases. This is the case both with and without optical filtering. To show the similarities between filtering at the transmitter and receiver, we compare the two in Figure 44 (b) with an eye diagram after transmission through 45-km of SMF. Filtering at the transmitter is nominally better than the one at the receiver. The position of the filter on the chirped spectrum has a significant effect on the Q factor, i.e. the filter slope should match the chirp of the laser, as shown in Figure 45, where we can see that asymmetric filtering (detuning from the spectrum peak by ~ 0.15-nm) can provide better performance, while centering the filter at the point of maximum output power is not an optimum position, as shown by the two eye diagrams inserted.
Figure 44 (a) Experimental results for optical filtering at the transmitter (after DML), two bias conditions are shown here (-0.9 V and -1.4 V); (b) Comparison of optical filtering at the transmitter and the receiver (bias=-1.4 V).

Figure 45 Comparison of optical spectrum and the corresponding eye diagrams for center filtering (maximum optical power) and side-band (asymmetric) filtering: ~ 20-GHz detuning from the carrier center is applied here.
The single link transmission results confirm the effectiveness of applying asymmetric narrowband optical filtering into directly modulated systems, and motivate us to find the feasibility of this approach in the long-haul transmission systems.

4.1.4 Long-haul dispersion managed link transmission

Besides single link transmission for metro and access networks, cost-effectiveness is still one of the major concerns for long-haul transmission systems, therefore the feasibility of direct modulation into such kind of systems is also attractive. However, dispersion management is required for such applications. In order to do so, we use a recirculating fiber-loop testbed to evaluate the transmission performance of direct modulation. Figure 46 shows the experimental setup of the recirculating loop testbed. The output power from the DML is ~5 dBm, with an extinction ratio of ~9 dB when driven with ~40 mA at a ~ -0.9 V bias. We place the DML into an 8-channel WDM system with 100-GHz (0.8 nm) spacing between channels. The DML was at ~1554.7 nm, the other seven channels are externally modulated using a LiNbO₃ electro-optic modulator and range from ~1552.2 to ~1557.8 nm. The external modulator and the DML are modulated at 9.95328 Gb/s with 2^{23}-1 PRBS, and are then decorrelated using a spool of fiber. The dispersion-managed recirculating loop consists of ~80-km SMF and ~12-km DCF (dispersion of -1348 ps/nm at 1550 nm). The input power per channel into the SMF and DCF
is fixed at 1.0 dBm and -4.0 dBm, respectively. Different short spools of SMF (0, ~2.5, and ~5.5-km) are inserted into the loop to vary the residual dispersion value – 0, ~2.5 and ~5.5-km SMF correspond to ~0.54 ps/nm/km, ~0 ps/nm/km and ~+0.60 ps/nm/km residual dispersion ($D_{res}$), respectively. A long-period-grating (LPG) is used to avoid noise accumulation and perform gain equalization along the link and an attenuator is used to balance the optical power without changing the OSNR. The receiver consists of a pre-amplifier, the same filter as used in single link transmission, and a PIN photodiode. The OSNR degradation along the link is ~21 dB after ~1400-km transmission. The output optical spectrum after ~1100-km transmission is shown in Figure 47 (a), and the relative location of optical filter to the modulated DML carrier center is shown in Fig 47 (b). Similarly, due to the intrinsic positive chirp of the DML, the carrier center is located at the negative sloping edge of the filter.

Figure 46 Experimental setup of long-haul DML transmission using a recirculating loop testbed.
Figure 47 (a) Optical spectrum of 8x10-Gb/s WDM system after ~ 1100-km transmission. (b) Relative location of optical filter (upper spectrum) to the DML carrier center (lower spectrum).

We took measurements for varying transmission distances so that we could determine the effects of asymmetric (sideband) or symmetric (center) filtering on the Q-factor. Figure 48 shows our results after ~940-km transmission. Figs. 48 (a) shows the output optical spectrum and the inserted eye diagram when using center filtering (no offset from the carrier), while Figure 48 (b) shows the output optical spectrum and the inserted eye diagram when using asymmetric filtering. In our experiment, center filtering cannot maintain error free transmission after only ~300-km transmission even in the negative residual dispersion case.
The overall transmission performance in terms of Q factor is shown in Figure 49.

As a typical one, Figure 49 (a) shows the Q-factor values for all the eight channels after ~1100-km transmission with ~0-ps/nm/km residual dispersion. The difference between the Q factor of the best externally modulated channel and the DML channel is ~0.8-dB, and all channels are error-free (Q > 15.6 dB). We believe that the major cause of Q degradation is OSNR degradation. Figure 49 (b) shows the Q-factor for the DML channel vs. transmission distance for varying values of residual dispersion – maximum error-free transmission distances are ~1400 km, ~1100 km, and ~580 km for residual link dispersion values of ~0.54 ps/nm/km, 0 ps/nm/km, and +0.60 ps/nm/km, respectively. Due to the positive chirp of our DML, the negative residual dispersion (-0.54 ps/nm/km) results in the best transmission performance.
Figure 49. Experimental results showing system performance improvement using asymmetric narrowband filtering. (a) Q-factor of eight channels after 1100-km transmission (under ~ 0-ps/nm residual dispersion). (b) Overall transmission performance (Q vs. distance) of DML under different residual dispersion values.

4.1.5 40-Gb/s Transmission Results

We further apply this technique into the 40-Gb/s directly modulated system using the same commercial DML. Our experimental setup is shown in Figure 50 (a). The DML is directly modulated at 40-Gb/s. The modulation voltage on the DML is ~2.8 V (peak-to-peak). The bias current is set to be ~110 mA – at such high current, the modulation bandwidth of the DML can be up to ~28 GHz. The output power is ~9 dBm. The extinction ratio of the laser is ~2.2 dB. After the DML, an optical filter with 3-dB bandwidth of 0.65-nm is used to improve the laser performance. The transmission spectrum and chromatic dispersion across the pass band are shown in Figure 50 (b). The transmission link is a single link using a spool of NDF without any optical amplification and dispersion compensation. The dispersion of NDF is
~1.7 ps/nm/km. The signal is detected using a PIN photodiode and de-multiplexed into four 10-Gb/s data streams and error analysis is performed at this data rate using an error detector.

Figure 50  (a) Experimental setup of 40-Gb/s DML transmission. (b) Chromatic dispersion and transmission spectrum of optical filter with 3-dB bandwidth of 0.65-nm.

In the back-to-back condition, without optical filtering, error-free reception cannot be achieved (BER floor at \(10^{-5}\)). However, when we apply an off-center 0.65-nm optical filter, the extinction ratio increases to \(~3\) dB, and error-free \((10^{-9})\) is easily obtained with a receiver sensitivity of \(~5.3\) dBm \(\sim-8\) dBm using external modulation in our setup). We use two spools of NDF for the transmission, 11-km and 25-km. The power penalties for 11-km and 25-km transmission are \(~3.5\) dB and 5 dB, respectively, as shown in Figure 51 (a). Typical eye diagrams are also shown in Figure 52 for comparison including the back-to-back cases (with and without filtering) and after 25-km transmission.
Filters of varying bandwidth are also used to find the limitations of filtering. Here we compare the effects of three filters with bandwidths of 0.30-nm, 0.65-nm and 1.4-nm. As shown in Figure 51 (b), a 0.3-nm filter can improve transmission performance, but still cannot reach error-free (an error floor at $\sim 10^{-7}$). This bandwidth might be too narrow for DML transmission although strong pre-filtering has been shown to improve performance in high spectrally-efficient transmission systems [97, 98]. A 1.4-nm filter is quite similar to the case without filtering with only slight improvement. These three filters all have similar Gaussian profiles and dispersion characteristics. We note that an optimized filter bandwidth (corresponding to a narrowed spectrum) with optimized detuning frequency (corresponding to a specific nonlinear phase response for FM to AM conversion) is essential to DML transmission. In our experiment, 0.65-nm turns out to be the best
bandwidth with a detuning frequency of ~10 GHz. The spectrum comparison is shown in Figure 53 (a) for the with and without filtering cases. We can see that the filter center is located at the left side of the center frequency of the carrier. This coincides with the reality that our DML has positive chirp. Another effect that we should also note is that the filter might need slight optimization under different link conditions. For example, as the link length increases from 11-km to 25-km, the best-case filter detuning frequency is slightly shifted to the left to, as shown in Figure 53 (b) with a detailed shift inserted.

![Figure 52](image1.png)  
**Figure 52** Eye diagrams of DML transmission (a) back-to-back without optical filtering; (b) back-to-back with optical filtering and (c) after 25-km NDF transmission using optical filtering

![Figure 53](image2.png)  
**Figure 53** Optical spectrum of filtering (a) comparison of spectrum with and without filtering; (b) Slightly detuning or optimization is necessary when the link length or dispersion varies (here the length changes from 11-km to 25-km).
To summarize this section, optical filtering can not only narrow the broadened spectrum due to direct modulation, but also provide the chirp management or FM to AM conversion via the interaction between the nonlinear phase response of the sloping edge of the optical filter, the laser chirp, and the link chromatic dispersion, therefore significantly improving system performance. We note that the narrowband optical filter can be integrated with the DML for further optimization or cost reduction.

4.2 Measurement of the Chirp Parameter of Electro-Optic Modulators By Comparing the Phase Between Two Side Bands

External electrooptic (EO) modulators have well-managed frequency chirp characteristics compared to directly-modulated semiconductor lasers and electroabsorption modulators. Chirp-free EO modulators, such as single-drive x-cut and dual-drive z-cut LiNbO$_3$ (LN) modulators, can eliminate the frequency chirping at the transmitters and repeaters of a fiber-optic network. However, in real-life fiber-optic networks, zero-chirp operation is not always desirable due to the chromatic dispersion of optical fibers [99, 100, 101]. Modulators with a negative chirp have been employed to reduce the transmission penalty due to chromatic dispersion [102]. Moreover, chirped modulation formats, e.g. chirped-return-to-zero (CRZ), have been demonstrated to be effective in ultra-long-haul
transmission systems with aggregate fiber capacities of up to Tbit/s [103]. Chirp on a signal can be realized using a phase modulator, or a single-drive or dual-drive z-cut LN intensity modulator [104]. Since the chirp parameter of optical modulators and the effective chirp of optical signals are key parameters in the design and evaluation of optical transmission systems, it is highly desirable that the chirp parameter of a modulator be obtained and verified using simple measurement techniques.

Previously reported methods for chirp measurement are based on: (i) the measurement of the sideband-to-carrier ratio of a chirped light spectra [105, 106]; (ii) the pulse width broadening after propagation through a dispersive fiber [107]; (iii) the frequency deviation due to chirp using a Mach-Zehnder interferometer as an optical discriminator [108]; (iv) measuring the induced Fabry-Perot fringes by the integration of the electro-absorption modulator with an active segment in a Fabry-Perot cavity [109]; and (v) the frequency-domain small signal frequency response through a dispersive fiber (typically 36 km to 401 km of single-mode fiber) found using a network analyzer [110].

In this section, we propose and demonstrate a simple technique for characterizing the chirp parameter of EO modulators without adding any optical fiber. This direct method applies a 10-GHz oscillator and a tunable fiber Fabry-Perot (FFP) filter. By selecting different parts (upper or lower) of the optical spectrum, or different
sideband signals created from the oscillator, we can measure the chirp parameter of an EO modulator by comparing the phase shift between the two optical sidebands [111]. Our technique can also distinguish the sign of the chirp, e.g. positive chirp will generate a positive phase shift (time delay) between the upper and lower frequency sidebands. In addition, this method can also be used to monitor the effective chirp of an optical signal along the transmission link.

![Diagram of experimental setup](image)

**Figure 54** Experimental setup for measuring the chirp parameter of optical modulators using partial side-band filtering and phase comparison.

The experimental setup is shown in Figure 54. As the signal chirping generated from an optical modulator produces a phase shift as the light intensity varies, we use a tunable optical filter to select the two distinct optical sidebands. The chirp (equivalent to a chromatic-dispersion effect) will induce a relative group delay, or phase shift, between the two sidebands. This phase difference can be readily detected using a phase detector or network analyzer.
The typical output amplitude of the electric field after a Mach-Zehnder EO modulator can be described as

\[ E_{\text{out}} = \frac{E_0}{2} e^{j\omega t} \cdot \sum_{k=-\infty}^{+\infty} ((j)^k J_k(m_1) + (j)^k J_k(m_2) \cdot e^{-j\varphi}) \cdot e^{j\omega t} \quad (3.6) \]

where \( E_0 \) is the amplitude of the electric field in the input waveguide, \( m_1 \) and \( m_2 \) are the optical phase modulation amplitudes, \( \omega_0 \) is the optical carrier frequency, \( \omega \) is the radio frequency (RF) modulation frequency, \( \varphi \) is a constant phase delay between the two interferometer arms, and \( J_k(x) \) is the k-th order Bessel function of the first kind. The chirp parameter \( \alpha \) can be related to the optical phase modulation amplitudes of the two modulator arms by [104]

\[ \alpha = \frac{m_1 + m_2}{m_1 - m_2}. \quad (3.7) \]

Based on Eq. 3.6, the RF modulation generates a series of sideband frequencies distributed at both sides of the carrier frequency. Our target here is to evaluate the RF signals generated by the interference of the carrier and the upper or lower sidebands. Assuming that an ideal optical bandpass filter (square pass band) is applied to select the carrier frequency and one of the side bands, we can calculate the output intensity at the photodetector. Because only the alternating current (AC) components are used in our measurement, we will only consider the frequency-dependent component in the output intensity. We designate the beat intensities generated by the carrier and upper/lower side band as \( I(\omega) \) and \( I(-\omega) \), respectively, leading to
\[ I(\pm \omega) = -2\sqrt{A^2 + B^2} \cos(\omega t \mp \theta), \quad (3.8) \]

where

\[ A = J_0(m_1)J_1(m_1) + J_0(m_2)J_1(m_2) + \left[ J_0(m_1)J_1(m_2) + J_1(m_1)J_0(m_2) \right] \cos \varphi, \quad (3.9) \]

\[ B = [J_0(m_2)J_1(m_1) - J_0(m_1)J_1(m_2)] \sin \varphi, \quad (3.10) \]

and

\[ \tan \theta = \frac{A}{B} = \frac{J_0(m_1)J_1(m_1) + J_0(m_2)J_1(m_2) + \left[ J_0(m_1)J_1(m_2) + J_1(m_1)J_0(m_2) \right] \cos \varphi}{[J_0(m_2)J_1(m_1) - J_0(m_1)J_1(m_2)] \sin \varphi}. \quad (3.11) \]

From Eqs. 3.8 and 3.11, the phase difference between the AC electrical signals corresponding to \( I(\pm \omega) \) is

\[ \Delta = \theta - (-\theta) = 2\theta \quad (3.12) \]

In order to link the phase delay angle \( \theta \) with the small signal chirp parameter \( \alpha \), we apply a small signal approximation to Eq. 3.11. Under this approximation, we have

\( J_0(x) \approx 1 \), and \( J_1(x) \approx x/2 \), thus

\[ \tan \theta \approx \frac{m_1 + m_2}{m_1 - m_2} \cot\left(\frac{\varphi}{2}\right) = \alpha \cot\left(\frac{\varphi}{2}\right). \quad (3.13) \]

For EO modulators biased at quadrature, \( \varphi=\pi/2 \). Thus, Eq. 3.13 provides a direct relationship between the small signal chirp parameter \( \alpha \) and the phase delay between \( I(\pm \omega) \) signals.
In measurements where the output signals are monitored by oscilloscopes rather than sensitive network analyzers, a large modulation depth is preferred to ensure a high signal-to-noise ratio. To evaluate $\alpha$ at large driving signal levels, we can combine Eqs. 3.7 and 3.11 to calculate $\alpha$ at different driving signal levels, as shown in Figure 55 for the phase delay as a function of chirp parameter $\alpha$ and Figure 56 for the phase delay as a function of modulation depth. From Figure 56, the phase delay half angle $\theta$ is close to a constant at small modulation depth and decreases slightly as the modulation depth increases.

Figure 55 Analytical result between phase difference and the chirp parameter.
Figure 56 Phase difference angle $\theta$ as a function of modulation depth for different chirp parameter values.

We experimentally characterize a LN modulator using the setup shown in Figure 54. In order to change the chirp value of the modulator, a dual-drive modulator driven by two signals divided from the same oscillator with different amplitudes is measured. A FFP filter with a FWHM bandwidth of approximately 8.0 GHz is used. The filter is detuned to generate 3.0-dB additional power loss, i.e. with approximately $\pm$ 4.7 GHz detuning. A phase sensitive detector is used to measure the phase shift as the optical filter is tuned over the signal spectrum.

Since the chirp is related to the bias voltage on the modulator as $\alpha_{\text{eff}} = \alpha \times \cot(\varphi/2)$ where $\varphi$ is proportional to the bias voltage (at quadrature bias, $\cot(\varphi/2) = 1$), first we bias the modulator at quadrature, i.e. the midpoint between the transmission null
and peak (in this case \(~-0.2\) V), and change the amplitudes of the sinusoidal signals on the two arms and measure the phase differences between the two sidebands for different chirp values according to the known driving voltages.

In order to increase accuracy, we further take into account the response of the optical filter since available optical filters typically have non-ideal response - they may cover more than one side band, or have a non-linear phase response. Using the characteristics of the FFP filter we used and considering above contributions, we get a new curve for the phase difference as a function of chirp parameter using computer simulation.

![Comparison of measurement results](image)

Figure 57 Comparison of measurement results (phase difference versus chirp parameter at the working bias voltage, i.e. tan\(\theta\)=1) and computer simulation results for a dual-drive electro-optic LiNbO\(_3\) modulator. Here an F-P filter with \(~8\) GHz bandwidth and 4.7 GHz detuning is used.
Figure 57 shows the comparison between measured phase differences versus the chirp parameters calculated from $\alpha = \frac{V_1 - V_2}{V_1 + V_2}$, where $V_1$ and $V_2$ are the corresponding applied voltages on the two arms, and the computer simulation results taking into account the filter response.

We then fixed the voltage on one arm and terminate the other arm, therefore the chirp parameter becomes 1 at the quadrature bias point. When we vary the bias voltage, the corresponding phase differences and output optical power change, as shown in Figure 58. We can see that the phase difference, i.e. effective chirp parameter, has a bias voltage dependence close to a tangential function.

Figure 58 The corresponding phase differences and output power at different bias voltages when the driving voltage applied to one arm of a dual-drive modulator.
4.3 Characterization of the Chirp Parameter of Electro-Optic Intensity Modulators Using Optical Spectrum Analysis

We have demonstrated a new approach for the measurement of the chirp parameter of EO modulator in the previous section [112]. In this section, we demonstrate a simple and straightforward method to measure the chirp parameter of EO modulators using optical spectrum analysis. The basic principle of this approach is: the sideband intensity distribution of a modulated signal depends on both the amplitude and phase modulation, as well as the interaction between them. By varying the DC bias voltage on the modulator, we can extract the amplitude and phase modulation contributions and obtain the chirp parameter from the intensity variation that appears the first sideband measured using an optical spectrum analyzer (OSA). This method only requires a sinusoidal signal generator, a DC power supply, and an OSA, all readily available instruments, while most of the previous methods need a certain length of fiber or other unstable equipments. Besides experimental simplicity, our method is accurate to within 5% due to the high-resolution and accurate intensity measurement capabilities of typical OSAs.

For a typical Mach-Zehnder EO interferometer, the output field $E_o$ under a sinusoidal modulation voltage can be expressed as
\[ E_o = A \exp(j\omega_0 t) \sum_{k=-\infty}^{\infty} (j)^k [J_k(m_1) + J_k(m_2) \exp(j\phi)] \exp(jk\omega_m t) \] (3.14)

where \( J_k(x) \) is the \( k \)th order Bessel function of the first kind, \( A \) is the amplitude of the input light, \( m_1 \) and \( m_2 \) are phase modulation amplitudes corresponding to interferometer arms 1 and 2 respectively, and \( \phi \) is the constant phase difference between the two interferometer arms. In EO materials, \( m_1 \) and \( m_2 \) are linearly proportional to the electric field in each interferometer arm. For a push-pull driving electrode structure, \( m_1 \) and \( m_2 \) have opposite signs. Equation 3.14 shows that the output electric field consists of a carrier component at \( \omega_0 \) and an infinite number of side bands at \( \omega_0 + k\omega_m \), where \( \omega_m \) is the external sinusoidal modulation signal frequency. The corresponding optical intensity of each frequency component is

\[
I(\omega_0 + k\omega) = |A|^2 \left| J_0(m_1) + J_0(m_2) \exp(-j\phi) \right|^2 = |A|^2 \left[ J_0^2(m_1) + J_0^2(m_2) + 2J_0(m_1)J_0(m_2) \cos \phi \right] 
\] (3.15a)

\[
I(\omega_0 + k\omega) = |A|^2 \left| J_0(m_1) + J_0(m_2) \exp(j\phi) \right|^2 = |A|^2 \left[ J_0^2(m_1) + J_0^2(m_2) + 2J_0(m_1)J_0(m_2) \cos \phi \right] 
\] (3.15b)

Thus, the ratio between the first two side bands (upper and lower) and the carrier intensity is

\[
\frac{I(\omega_0 \pm \omega)}{I(\omega_0)} = \frac{J_1^2(m_1) + J_1^2(m_2) + 2J_1(m_1)J_1(m_2) \cos \phi}{J_0^2(m_1) + J_0^2(m_2) + 2J_0(m_1)J_0(m_2) \cos \phi} 
\] (3.16)

where the factor \( k \) is the sideband number. Since \( \phi \) is related to the control bias voltage, the bias voltage can be varied such that \( \cos(\phi) \) becomes +1 or -1, and the maximum and minimum intensity for the first (upper and lower) sidebands (note \( m_1 \) and \( m_2 \) have opposite signs) can be determined to be
\[ I(\omega_0 \pm \omega)_{\text{max}} = \left| J_1(m_1) - J_1(m_2) \right|^2 |A|^2 \]  

(3.17a)

and

\[ I(\omega_0 \pm \omega)_{\text{min}} = \left| J_1(m_1) + J_1(m_2) \right|^2 |A|^2 \]  

(3.17b)

Using the relationship between \( m_1, m_2 \), and the chirp parameter \( \alpha_c \) at quadrature, we easily find the chirp parameter from the maximum and minimum intensities of the first sideband

\[ \alpha_c = \frac{m_1 + m_2}{m_1 - m_2} = \sqrt{\frac{I_{\text{min}}}{I_{\text{max}}}} \]  

(3.18a)

or

\[ \alpha_c = \sqrt{10^{\frac{I_{\text{min}} - I_{\text{max}}}{10}}} \]  

(3.18b)

where \( I_{\text{min}} \) and \( I_{\text{max}} \) are measured in Watt, and \( I'_{\text{min}} \) and \( I'_{\text{max}} \) are measured in dBm.

Note that we use the small signal approximation for \( J_1(m) \), i.e., \( J_1(m) = m/2 \).

Therefore, in order to evaluate the chirp parameter of the modulator, our self-referenced approach measures the maximum and minimum intensities of the first sideband when the phase bias \( \phi \) in Eq. 15 is adjusted by applying a DC voltage \( V_{\text{bias}} \). As shown in Figure 59, a 10-GHz oscillator generates a sinusoidal signal \( (V_{\text{RF}}) \) that is applied to the EO modulator, and sideband intensities at different frequencies are measured with an OSA. As the DC bias varies, the output spectrum
changes accordingly. From the variation of the intensity of either the upper or lower first sideband, we calculate the chirp parameter using Eqns. 3.18a or 3.18b.

Figure 59 Conceptual diagram of chirp measurement using optical spectrum analysis: optical intensity variation of the first sideband of the signal corresponds to the chirp value. In the case of a dual-drive modulator, the 10-GHz RF signal is applied on both arms with an appropriate time delay.

Using optical spectrum analysis, the high frequency (>10 GHz) half-wave voltage $V_{\pi}$ can also be readily obtained by comparing the maximum intensity difference between the carrier and first sideband when the bias voltage $V_{bias}$ is varied, as described by

$$V_{\pi} = V_{RF} \times \frac{\pi}{4} \sqrt{\frac{I(\omega_{0})_{\text{max}}}{I(\omega_{0} \pm \omega_{m})_{\text{max}}}}$$

(3.19)

where the sinusoidal RF voltage $V_{RF}$ can be accurately measured by an RF power meter or a broadband oscilloscope.

In order to verify the accuracy of our method, we choose a 16-Gb/s dual-drive modulator and measure the chirp parameter using the setup shown in Figure 59. We
fix the driving voltage of the 10-GHz signal ($V_{\text{RF}}$) on one of the MZ arms ($V_1 \sim 2$ V), and vary the voltage $V_2$ on the other arm by using a tunable electrical attenuator, such as a tunable delay line to match the phase between the two arms. Different chirp values are obtained according to the equation $(V_1 - V_2)/(V_1 + V_2)$ under the small signal approximation [113]. The resolution of the OSA is set to 0.01 nm. However, as the resolution decreases, such as 0.05 nm, the measurement accuracy decreases. A 5-GHz oscillator may also be used with 0.01-nm OSA resolution. As shown in Figs. 60 and 61, we can calculate the chirp value from the intensity variation of the first side band during OSA scans. Figure 60 shows the agreement between our method and direct calculation.

![Figure 60](image)

**Figure 60** Comparison between chirp measurement results using our method and calculated from $(V_1 - V_2)/(V_1 + V_2)$ method. Here the chirp value is varied by using a dual-drive modulator and changing the driving voltage on one arm.
Figs. 61 (a)-(c) show several typical optical spectra with different chirp values. For a chirp parameter of 0, the intensity of the first side band can change by over 25 dB. Additionally, for a chirp parameter of 1.0 with one of the arms terminated, the intensity of the first side band is almost constant as the bias voltage changes. For values between 0 and 1, the chirp can be easily calculated using Eqs. 3.18.

Another key issue is that the driving signal amplitude will affect the measurement results. Data sheets of commercial EO modulators give the chirp values (between 0 and 1) at a given driving voltage. As we are using a small signal approximation in our formula (Eqs. 18), slight variations can arise as the driving voltage changes. In order to measure this effect, we use a commercial 40-Gb/s single drive z-cut LiNbO₃ modulator and apply different voltages to the modulator [114]. The optical spectrum varies slightly as the driving voltages changes, and we can obtain the relationship between the measured chirp parameter and the driving signal.
amplitude (Figure 62). As the signal increases, the measured chirp value decreases slightly. At a specified voltage, our measured chirp value agrees well with the modulator data sheet.

![Figure 62 Effect of the signal amplitude on the measurement results of chirp parameter using a commercial modulator with a typical spectrum inserted as the chirp value ~ 0.75.](image)

In our experiments, the accuracy of our method is estimated to be <5% and depends on the accuracy of our OSA, laser stability, and precise voltage adjustment.
4.4 Performance Optimization of RZ Data Format in WDM Systems Using Tunable Pulse-width Management at the Transmitter

In ≥10-Gbit/s wavelength-division-multiplexed (WDM) transmission systems, channels experience a confluence of channel-degrading effects that relate to chromatic dispersion and fiber nonlinearities [13, 14]. These high-speed, long-distance, or high-spectral-efficiency systems must be designed very carefully such that the data integrity will remain high at the receiver. Rigorous standards are required for deployment: (i) extensive fiber-link measurements are taken before system design, (ii) equipment must be manufactured to exact values to account for various deleterious fiber effects, and (iii) the fiber plant cannot change after system deployment for fear of system outage. Moreover, any system upgrades will require the replacement of much of the existing terminal equipment since management of the fiber effects will change significantly. All these factors incur large operational costs for system providers as well as severe constraints on the network operators. On the other hand, due to the dynamic nature of today’s reconfigurable networks, link conditions and networking performance (such as chromatic dispersion, fiber nonlinearities, polarization effects, etc.) may vary from time to time due to path changes or environmental perturbations as data traffic travels around.
A laudable goal is to enable tunability and flexibility in the terminal equipment to enable proper system performance even under conditions that may vary over time. One important scenario is to provide tunability in the data pulse width that is launched into a system for return-to-zero (RZ) systems. Pulse compression using phase modulators and dispersive fibers has been discussed in the earlier work [115], and recent simulations have shown that the performance of a link can vary significantly depending on the pulse width, even for small changes in fiber characteristics [116, 117]. Such pulse-width management can be accomplished by tuning the phase-modulator chirp as well as tuning the dispersion value of a dispersive element. Appropriate combination of chirp and dispersion values can compress or expand the data pulse. To date, although some experimental results show the performance of systems employing different relative pulse widths [118, 119], no demonstration has shown practical tunable pulse-width management and the performance improvement by deploying such kind of management. Moreover, another potential valuable issue would be to provide pulse-width management for several WDM channels simultaneously.

In this section, we demonstrate both single channel and 4x10-Gb/s WDM system optimization using tunable pulse width management of the RZ format. RZ pulse widths from 50 ps to 10 ps can be obtained by tuning a phase modulator and a tunable dispersion element at the transmitter. We show that even under different
environmental conditions or different system configurations, system performance (both the single channel and WDM systems) can be optimized by varying the pulse width. For example, as the link dispersion varies by only ~4%, for 10-Gb/s single channel transmission with 50-ps RZ pulses, the transmission distance with ~5 dB power penalty varies between 1200 km and 600 km. Using pulse management, the 5-dB distance can be almost doubled to 2400 km and 1000 km, with the same residual dispersion. For a 4x10-Gb/s WDM system with 0.8-nm channel spacing, we realized doubled transmission distance for specific residual dispersion. Tunable pulse width management can significantly enhance the robustness of WDM RZ transmission systems to nonlinearities and increase spectral efficiency [120].

4.4.1 Tunable Pulse-Width Generation

Pulse compression using electro-optic phase modulators and a dispersive medium (such as optical fiber) has been previously reported [115, 120], as have some other methods [121]. In general, the complex optical electric field of a chirped pulse can be expressed as [115]

$$\cos\left[\frac{\pi}{2}\sin(\Omega t)\right]e^{i(\omega t+\frac{A}{2}\sin(2\Omega t))} \tag{3.20}$$

where A is the peak-to-peak amplitude of the phase modulation, \(\omega\) is the angular frequency of the optical carrier, and \(\Omega\) is the angular frequency of driving signal. A Fourier transform yields the discrete frequency spectrum of the pulse. The effect of
propagating through a dispersive medium can be obtained by multiplying each spectral component at angular frequency \( \omega \pm n \Omega \) by \( \exp(-i \lambda^2 D L n^2 \Omega^2 / 4 \pi c) \) where \( \lambda \) is the wavelength of the carrier, \( D \) is the dispersion of the fiber, and \( L \) is the fiber length, \( c \) is the speed of light in vacuum [115]. Using inverse Fourier transforms, we can get the corresponding pulse width considering the full-width at half maximum (FWHM). Although this is well known, in order to apply it into our experiments, we use computer simulations to get the narrowest pulse width (pure RZ signal) at different chirp values and dispersion values, as shown in Figure 63. Here we assume that the input RZ signal is at a 10-Gb/s data rate with a duty-cycle of 0.5 (i.e. 50-ps pulse). Figure 63 (a) only shows four chirp values: 0, \( \pi/2 \), \( \pi \), and \( 2\pi \), while Figure 63 (b) shows the typical points that we used in our experiments (pulse width of 10-ps, 25-ps and 35-ps).

![Figure 63 Tunable pulse-width generation using chromatic dispersion chirped signal.](image-url)
4.4.2 Experimental Setup

The experimental setup for tunable pulse width RZ transmission is shown in Figure 64. Four WDM channels located at 1553.4, 1554.2, 1555.0 and 1555.8 nm are multiplexed and RZ modulated using two cascaded electro-optic modulators. The first amplitude modulator is driven by $2^{23}-1$ PRBS at 10-Gb/s data rate. An RZ signal with 50-ps pulse width is obtained after the second amplitude modulator, which is driven by the clock signal. A phase modulator driven by the clock followed by a multi-channel tunable dispersion element is used to generate tunable pulse widths. Different tunable dispersion elements can be used, such as [122, 123, 124, 125]. The corresponding pulse compression (from 50 ps to as low as 10 ps) can be realized using different amounts of chirp and dispersion as described in the previous section. The amplitude of the driving signal to the phase modulator can be up to 8 V, which is the half-wave voltage of the modulator. The narrowest pulse-width in our experiments is ~ 10 ps, achieved when we apply 8 V amplitude to the phase modulator (the corresponding chirp value is $\pi$) followed by ~ 60 ps/nm dispersion. The dispersion of the mechanically stretched sampled nonlinearly-chirped FBG can be tuned from 600 to 1900 ps/nm for the four channels, with the corresponding dispersion vs. wavelength curve shown in Figure 65.
Figure 64 Experimental setup for tunable pulse width RZ transmission using a recirculating loop testbed. AM: amplitude modulator; PM: phase modulator; GEQ: gain equalizer; OF: optical filter.

Figure 65 Dispersion vs. wavelength of the sampled nonlinearly-chirped FBG

A spool of SMF is used to offset the grating to get the desired positive dispersion from 60 to 160 ps/nm. For single channel transmission, another single-channel nonlinearly-chirped FGB is used with a tuning range between 100 and 500 ps/nm, as well as different spool of SMF to offset this dispersion, too. The group delay ripples of our gratings are $\sim \pm 10$ ps.
After pulse shaping, the signal is transmitted through a recirculating loop testbed, which consists of three Erbium-doped fiber amplifiers (EDFAs) working in the saturated regime, 78-km SMF and 12-km DCF. A spool of SMF with different lengths from 0 to 4.4 km is used to change the residual dispersion while maintaining the optical signal-to-noise-ratio (OSNR) of the link using an optical attenuator. A long-period-grating (LPG) works as gain equalizer and reduces amplified-spontaneous-emission (ASE) noise. A typical optical spectrum after 8 round trips through this testbed (i.e. a given distance) is shown in Figure 66. An optical pre-amplifier before the receiver is used to increase sensitivity after long-distance transmission. Power penalties, including OSNR degradation, are measured by comparing the receiver sensitivity at $10^{-9}$ bit-error-rate (BER) with the back-to-back one (here $\sim -34$ dBm for the RZ format).

![Figure 66 Optical spectrum after 8-loop transmission](image_url)

Figure 66 Optical spectrum after 8-loop transmission
4.4.3 Experimental Results and Discussion

We first compare the transmission performance for the single-channel case by varying the pulse width. Figure 67 (a) shows the measured power penalties after 600-km transmission for different residual dispersion values. The input power is set to ~ 6 dBm for the 78-km SMF and to ~-1 dBm for the DCF. Three typical residual dispersion values have been used by changing the length of the small spool of SMF: 0.4 ps/nm/km, 0.08 ps/nm/km and –0.2 ps/nm/km, which is equivalent to ~ 4% link dispersion variation. In terms of the transmission distances at ~ 5 dB power penalty (Figure 67b), the transmission distance for 50-ps RZ pulses is reduced from 1200 km to 600 km as the residual dispersion varies, while the distance can be extended to between 2400 km and 1000 km with the same residual dispersion using optimized pulse widths (35 ps and 25 ps in our experiment). For single channel transmission, shorter pulse widths are more sensitive to dispersion variation, consequently the optimal pulse width for different dispersion values is ~ 25-35 ps, which is close to our simulation results.
The experimental results of 4x10-Gb/s WDM transmission are shown in Figure 68(a)-(c). The channel spacing is 0.8 nm and the sampled grating is tuned via stretching for all four channels simultaneously, thus the same pulse width is obtained for all the channels. The channels are decorrelated through about 10-km of dispersive fiber. The optical power inputs for 78-km SMF and DCF are set to ~4 dBm and ~2 dBm per channel, respectively. The transmission performance of a typical channel (1555.0 nm) is measured. For the residual dispersion value of 0.4-ps/nm/km, the transmission distances with ~4 dB power penalty are 1900, 1600, and 800 km for pulse widths of 50, 35, and 25 ps, respectively. As the residual dispersion value gets close to zero (0.08-ps/nm/km), the distances change to 800, 1300, and 1200 km for the above listed pulse widths. Although the experimental results under zero and negative residual dispersion are a little different from the
simulation results, this may be due to some special link conditions, such as different fiber nonlinearities, EDFA saturation, etc.

Since fiber nonlinearities, such as self-phase-modulation (SPM) and cross-phase-modulation (XPM), are key issues in single channel and WDM systems [126], managing pulse width at the transmitter enables flexibility in terms of system optimization.

To summarize this section, we demonstrate performance optimization for RZ data transmission using tunable pulse width management at the transmitter. Tuning the driving voltage on a phase modulator and the dispersion value of a tunable dispersion element generates the tunable pulse width. Moreover, it is easy to alter this RZ transmission to generate the chirped RZ (CRZ) format [103], which can

![Graphs showing performance optimization for different pulse widths](image-url)
also converge to the dispersion managed soliton format [127]. In addition, the stability (compared to mode-locked lasers) and tunability can enhance system robustness and increase the system power margin for both single channel and WDM systems. Although we performed the demonstration using a 10-Gb/s data rate, if the half-wave voltage of phase modulators can be reduced further, dynamic pulse compression down to several picoseconds is possible. This will ensure high spectral efficiency and high capacity transmission at even higher data rates (>40-Gb/s).

### 4.5 Uniformly Distributed States of Polarization on the Poincarè Sphere Using an Improved Polarization Scrambling Scheme

As mentioned times in this thesis, the once-disregarded effects of lightwave polarization on a system can cause significant network impairment as optical fiber transmission systems begin to employ higher data rates (>10-Gb/s/channel). In particular, PMD and PDL [42, 43, 128], as well as PDG [43] have received recent attention. In addition, the random nature of polarization effects is also a major problem for polarization-sensitive instrumentation. Scrambling the states of polarization (SOPs) of the signal has been shown to be a valuable technique that
can facilitate the compensation or reduction of polarization-related impairments [43, 52-54, 129-132], or the reduction of measurement uncertainty.

Typically, polarization scrambling schemes are based on high-speed polarization modulators or relatively low-speed polarization controllers, for example, LiNbO$_3$ devices (>1 MHz bandwidth) [50, 133] or fiber-squeezer based polarization controllers (up to 100 kHz bandwidth). LiNbO$_3$ polarization modulators have been used as bit-synchronous polarization scramblers to reduce PDG effects in undersea transmission systems where the scrambler output SOP is modulated along a great circle on the Poincaré sphere at the bit rate [48]. When high-speed polarization scrambling is not required, sets of polarization controllers are used as scramblers. Scramblers with different scrambling speeds and different output SOP distributions (corresponding to different reduced degrees of polarization (DOPs)) may have different effects on system performance. Moreover, it is often necessary that the scrambler output SOP is distributed, preferably uniformly, on the entire Poincaré sphere for the purposes of PMD and/or PDL monitoring and compensation [52-54, 130-132]. In this section, we concentrate on this type of uniformly-distributed polarization scrambling. Specifically, we discuss fiber squeezer-based polarization controllers, which have the advantages of low cost and ultra-low intrinsic PDL (i.e. negligible intrinsic signal degradation).
One common fiber-squeezer-based polarization-scrambling scheme uses varying control voltages (each a different waveform) applied to each of the three sections of the polarization controller, known as the three-function approach. These control functions vary in both frequency and peak-to-peak voltage. Although this method does move the output SOP over the Poincarè sphere, it suffers from uncertainty (it is randomly non-repeatable) from measurement to measurement, and it is difficult to uniformly distribute the output SOP over the sphere. To obtain the most efficient PMD and/or PDL monitoring and compensation, it is desired that the output (scrambled) SOPs are uniformly distributed and repeatable.

Here we propose a periodic random polarization scrambling scheme that can generate uniformly distributed output SOPs on the Poincarè sphere for any arbitrary input polarization state using a three-section fiber-squeezer based polarization controller.

### 4.5.1 Uniformly Random Polarization Scrambling

Typical fiber-squeezer-based polarization controllers consist of multiple fiber squeezers oriented 45° with respect to each other (Figure 69(a)). Each fiber squeezer is driven by an applied voltage signal on the fast electronic actuator. Squeezing the optical fiber produces a linear birefringence in the fiber and thus alters the state of polarization of a light signal passing through it. i.e. induces phase...
retardation. For a given input polarization state, the first section can make the polarization rotate around a fixed axis on the Poincare sphere (named $s_1$), the second section will rotate around another axis that is perpendicular to the first one (named $s_2$), as shown in Figure 69(b). The third section always has the same characteristics as the first section with a little perturbation to facilitate the realization of input polarization independent.

![Figure 69](image)

Figure 69 (a) Configuration of 3-section fiber squeezer polarization controller and (b) the corresponding polarization state traces on the Poincare sphere as driving voltages vary.

For a three-section fiber squeezer scrambler, the transfer matrices for the three sections are [134, 135]

$$
\begin{bmatrix}
\exp(i\frac{\theta_1}{2}) & 0 \\
0 & \exp(-i\frac{\theta_1}{2})
\end{bmatrix},
\begin{bmatrix}
\cos\frac{\theta_2}{2} & i\sin\frac{\theta_2}{2} \\
\sin\frac{\theta_2}{2} & \cos\frac{\theta_2}{2}
\end{bmatrix},
\text{and}
\begin{bmatrix}
\exp(i\frac{\theta_3}{2}) & 0 \\
0 & \exp(-i\frac{\theta_3}{2})
\end{bmatrix}
$$

(3.21)

where $\theta_1$, $\theta_2$, and $\theta_3$ are the corresponding phase retardations of each section, which are proportional to the control voltage on that section.
In order to realize uniformly distributed SOPs for any arbitrary unknown input SOP, we analyze the polarization transfer matrix and derive a random polarization transformation matrix as \[41\] (also in section 3.1)

\[
P = \begin{bmatrix}
\sqrt{\gamma} \exp(i\theta) & -\sqrt{1-\gamma} \exp(-i\phi) \\
\sqrt{1-\gamma} \exp(i\phi) & \sqrt{\gamma} \exp(-i\theta)
\end{bmatrix}
\]

(3.22)

where \(0 \leq \gamma \leq 1\), \(-\pi \leq \theta \leq \pi\), and \(-\pi \leq \phi \leq \pi\) are statistically independent and uniformly-distributed random variables. Given an arbitrary 2×2 unitary matrix \(U\), it can be shown that the matrix product \(UP\) or \(PU\) has identical distribution as \(P\) (see the Appendix of [41] for a mathematical proof). Therefore, the random transformation \(P\) scatters the SOP uniformly on the Poincaré sphere for any input SOP.

If we multiply the matrices of the three-section squeezer, we can get the product as a new matrix:

\[
Q = \begin{bmatrix}
\cos \frac{\theta_2}{2} \exp(i \frac{\theta_1 + \theta_3}{2}) & i \sin \frac{\theta_2}{2} \exp(-i \frac{\theta_1 - \theta_3}{2}) \\
i \sin \frac{\theta_2}{2} \exp(i \frac{\theta_1 - \theta_3}{2}) & \cos \frac{\theta_2}{2} \exp(-i \frac{\theta_1 + \theta_3}{2})
\end{bmatrix}
\]

Comparing this matrix \(Q\) with the matrix \(P\), we find that there are some relationships between \(\theta_1, \theta_2, \theta_3\) in matrix \(Q\) and \(\gamma, \theta, \phi\) in matrix \(P\):
$\frac{\theta_1 + \theta_3}{2} \leftrightarrow \theta, \quad \frac{\theta_1 - \theta_3 - \pi}{2} \leftrightarrow \phi,$
and
$\cos^2 \frac{\theta_2}{2} = \frac{1 + \cos(\theta_2)}{2} \leftrightarrow \gamma \quad (3.23)$

In order to obtain uniformly distributed SOPs, we can distribute $\theta_1$ and $\theta_3$ uniformly between $[-\pi, +\pi]$, while distributing $\cos \theta_2$ uniformly between $[-1, +1]$.

The phase retardations of fiber squeezers are proportional to the applied voltages, thus the required distributions of phase retardation can be mapped to the corresponding driving voltages on each section, e.g. $V_1$ (driving voltage on the first section) can be uniformly distributed between $[-V_\pi, +V_\pi]$. According to this scenario, a series of random control voltages are generated according to this distribution and we periodically apply them to the polarization scrambler to perform uniformly random polarization scrambling. While it is possible to also distribute $\theta_2$ between $[-\pi, +\pi]$ instead of using our proposed approach of distributing $\sin \theta_2$ between $[-1, +1]$, the output uniformity is greatly decreased in the process - we call this approach “pure random” polarization scrambling.

4.5.2 Simulation results

We simulate the system to compare the output SOPs using the proposed approach - uniformly random polarization scrambling, with the traditional three-function approach, as well as the pure random approach.
Figure 70 compares the output degree-of-polarization (DOP) as the sample number increases for our periodic random scrambling and the three-function approaches. Lower DOP can reduce the effects of PDG and the uncertainty of the output polarization variation. Here the DOP is defined as

\[
DOP = \sqrt{\langle S_1 \rangle^2 + \langle S_2 \rangle^2 + \langle S_3 \rangle^2 / S_0}
\]  

(3.24)

where \(S_0, S_1, S_2, \) and \(S_3\) are the Stokes parameters of the output SOPs, and \(\langle S_i \rangle\) refers to the time averages of Stokes parameters. As the number of samples rises, the DOP with periodic uniform polarization scrambling nearly reaches 0%, while the three-function scrambler approaches a finite value. In this simulation, we used sinusoidal functions with frequencies of 31.1, 67.7, and 107.3 kHz with a sampling rate of 64x107.3 kHz. Note these frequencies are chosen according to the resonance frequencies of the commercial PZT components we used in our experiments.
Working at these frequencies can significantly reduce the power consumption. If other frequencies are used, the DOP may be further reduced [133].

One may say that the DOP can be further reduced using the N-function driving method if more sections of squeezers are used. This is true as long as we choose the driving frequencies wisely. However, this won’t help to improve the generated SOP uniformity. In addition, the reduced DOP also depends on the input SOP, i.e. the output SOP may vary by 10% for different input polarization states.

In addition, the uniformity of the output SOPs using the two approaches are compared in Figs. 71 and 72. In Figure 71, we compare the output SOPs for 1000 samples using the two approaches. Periodic uniformly random scrambling (Figure 71(a)) covers the entire sphere well, while three-function scrambling (Figure 71(b)) has some “holes” after scrambling. The same frequencies of sinusoidal signals are used here.
Figure 71 Output SOPs of 1000 samples using (a) periodic uniformly random polarization scrambling and (b) three-function polarization scrambling showing “holes” on the sphere.

In a three-dimensional space, if the output points on a surface of the sphere are uniformly distributed, the projections of these points on any major axes should also be uniformly distributed, i.e. if output SOPs on the sphere after scrambling are uniformly distributed, one of the necessary requirements is that the average optical power after an arbitrary polarizer should be constant. Instead of gathering the optical power after the polarizer, we look into the projections on the three Stokes axes. Figure 72 shows the uniformity of the Stokes parameters using three approaches: (a) is three-function scrambling, (b) is the “pure random” scrambling approach with all three phase retardations uniformly distributed between [-π, +π], and (c) is our proposed approach, distributing cosθ2 between [-1, +1]. We can see that neither (a) nor (b) provide a uniform distribution.
4.5.3 Measurement results

In order to experimentally test the output uniformity, we design the control software to generate a series of voltages and apply them to a fiber-squeezer-based polarization controller, and then measure the corresponding output SOPs using a commercial polarization analyzer. The input polarization state is set to an arbitrary state. In fact, we have also applied this scrambling scheme to generate the required uniform distribution [136] where the polarization coupling is totally random. For ~1000 control samples (i.e. 1000 voltage series), we use computer to obtain the
output SOPs from the polarization analyzer one by one and then plot them on the
Poincarè sphere as shown in Figure 73(a), while 73(b) shows the uniformity of the
Stokes parameters. As mentioned in the previous section, we know that this is only
a necessary requirement for the uniformity of SOP distribution.

![Figure 73](image)

**Figure 73** (a) Measured output SOPs of ~1000 samples on the Poincarè sphere.
(b) Distribution of the Stokes parameters.

There is a practical issue for this application. For each section of commercial fiber-
squeezer based polarization controllers, there is always some pre-loading (intrinsic
pressure) on the squeezer even without electrical driving signals on it, thus
introducing a “biased” term on the transfer matrices. Since $\theta_1$ and $\theta_3$ are uniformly distributed and proportional to the driving voltages, these two phase-retardations won’t effect the output distribution. However, in order to get rid of the contribution of a “biased” $\theta_2$, we need to find out an intrinsic driving voltage under which the second section is “transparent” to the first and the third section. To do that, first we align the input SOP with the eigen polarization state of the first section, and then tune the voltages on the second and the third section. As long as the SOP after the second section (under a certain voltage) becomes the eigen polarization state to the third one as well, then the voltage on the second can be used as the mid-point voltage for the generation of series of applied voltages. The measurement results confirm that our approach results in the most uniformly-distributed SOPs after we took into account these practical issues.

Another practical issue is the driving bandwidth. For the fiber squeezer polarization controller, the sinusoidal drive has a well-defined bandwidth up to tens of kHz according to the characteristics (especially the resonance frequencies) of the PZT. In principle, the bandwidth of proposed polarization scrambling can be hundreds of kHz, however, due to the rise time of the PZT (e.g. ~ 100 $\mu$s at 60-V voltage) and the mechanical loading response, the bandwidth is also at the order of tens of kHz. In our experiments, > 20-kHz update frequency has been applied successfully.
4.5.4 System applications

As mentioned in the introduction, polarization related impairments could be difficult to deal with due to their dynamic and random nature. This approach, however, is one of the key techniques that can be applied to overcome or mitigate these impairments in transmission systems, or to measure device polarization characteristics (e.g. PDL of optical devices), or to reduce the uncertainty when measuring polarization-dependent system parameters.

(a). PDG mitigation: Because the Erbium doping is not perfect in optical fiber amplifiers, the PDG-induced polarization-hole-burning (PHB) effects along the link may cause system fluctuation. This degradation may be suppressed by scrambling the SOP [43, 129]. Bit synchronized or low frequency polarization scrambling can reduce this effect [129, 48]. In this case, uniformly-distributed polarization scrambling can be used but not necessary.

(b). PDL characterization and compensation: Most of the optical components used in fiber communication systems have some sort of PDL, defined as the maximum loss difference for all the input polarization states or $I_{\text{max}} - I_{\text{min}}$. PDL can induce fluctuations in the optical signal-to-noise-ratio (OSNR). Since the polarization coupling between cascaded PDL sources along the link changes randomly, system
performance may vary from time to time [43]. In order to monitor and compensate
the PDL along the transmission link, a fast polarization scrambler should be placed
after the transmitter, then the PDL value can be monitored from the power
fluctuation induced by PDL of the components or optical modules (e.g. EDFAs)
[54], the feedback signal controls a polarization controller (like in PMD
compensation) and a variable PDL source to minimize the power fluctuation
(Figure 74(a)). This can also be used to perform fast and accurate PDL
characterization of fiber-optic devices in a manufacturing environment. Details
about this application will be discussed in the later section.

(c). PMD monitoring and compensation: PMD arises along the transmission fiber
due to the different propagation speed of two orthogonal polarization states [42,
128]. A typical first-order PMD compensator (Figure 74(b)) is comprised of a
dynamic polarization controller followed by a fixed or variable differential-group-
delay (DGD) element. As PMD along the link increases, the DOP of the signal will
decrease, thus by monitoring the DOP value at the end of the link, we can get the
effective PMD information. However, since the DOP depends on the signal SOP at
the link input, the monitored DOP can not provide the exact instantaneous DGD
value of the link. This in turn may cause feedback fading (i.e. a fall into a local
minimum) during polarization tracking. In this case, we can apply polarization
scrambling at the beginning of the link and by monitoring the instantaneous DOP
information, we can get the instantaneous DGD value of the link (or the principal state of polarization (PSP) of the link) [52, 53, 130-132]. Instead of feedback control, we can preset the DGD value of the variable DGD compensator based on the monitored instantaneous DGD value, and then use another performance monitor that correlates PMD with some parameter besides RF spectrum analysis and DOP information (i.e. eye opening, etc.) as feedback to track the polarization controller. This configuration may seem a little complicated, but it can decouple the monitored information and avoid the local minimum problem.

![Diagram](image)

**Figure 74** System applications of uniformly-distributed polarization scrambling: (a) PDL monitoring and compensation (b) PMD monitoring and compensation.
In addition to the system applications, polarization scrambling can be used to eliminate an instrument’s polarization sensitivity. Some optical instruments, such as diffraction-grating-based optical spectrum analyzers, are sensitive to the SOP of the input light. Scrambling the input polarization can remove the measurement uncertainties caused by polarization sensitivity.

4.6 Demonstration of In-line Monitoring and Compensation of Polarization Dependent Loss for Multiple Channels

Because many optical in-line components (e.g., couplers, switches, EDFAs, and isolators) produce non-negligible PDL, such PDL induced effects can become pronounced and produce more complex system related effects than originally assumed [23, 44-47, 137].

Some of the deleterious effects induced by PDL may include: (i) variation in the optical power and signal-to-noise ratio (SNR) of each WDM channel, and (ii) mutual interaction between PMD and PDL that tends to broaden the distribution of system power penalties and invalidate some PMD compensation [44-47, 137]. Furthermore, the aggregate PDL of a combination of several PDL-generating devices is wavelength dependent in the presence of PMD which randomizes the
signal's polarization state as an uncorrelated wavelength function, thus the effects of PDL for many WDM channels will be uncorrelated.

A practical scheme of fast PDL monitoring and dynamic compensation would be of significant value for high-performance systems. However, two specific challenges are: (i) although the PDL value of a single component is fixed, the aggregate PDL value of several in-line components or an optical module is a time-varying function due to the random signal polarization state between each component, and (ii) <millisecond-time-scale erbium-doped fiber amplifier (EDFA) output power saturation [138] may distort the measurement of PDL in that polarization-induced power fluctuations are not linearly transmitted through the EDFA.

In this section, we demonstrate a scheme for in-line PDL monitoring and compensation. In order to avoid the influence of EDFA output power saturation, the monitoring is accomplished by using >20-kHz polarization scrambling of either: (i) the data wavelength channel itself, or (ii) a wavelength that is ancillary to the collection of WDM channels. The corresponding PDL compensation is accomplished by using the monitored PDL value as a control signal to vary a tunable PDL module and minimize the power fluctuations. We perform PDL monitoring and compensation periodically along an 800-km link for four 10-Gbit/s
WDM signals. In the presence of 14 ps of average PMD, the PDL compensator reduces the 2% power penalty distribution tail from 6.5 dB to < 2.0 dB.

### 4.6.1 Concept of In-Line PDL Monitoring and Compensation

In order to monitor the PDL along a link and avoid the influence of EDFA power saturation, we use kHz-rate polarization scrambling, or repeated scanning, over a series (typically 100) of uncorrelated states-of-polarization (SOP) at the starting point of the link (see Figure 75). The instantaneous PDL value of a module that contains several PDL-generating devices is obtained from the root-mean-square (RMS) variation of the photo-detected signal power induced by PDL. Figure 76 shows the RMS power variation of the monitored signal after passing through a 0.8-dB PDL component as a function of the polarization scrambling frequency. For our experiment, the influence of EDFA transients is suppressed for a scrambling frequency >2 kHz, and the results are consistent with measurements taken for a PDL-generating module that has no EDFAs (i.e. passive module).

![Figure 75. Schematic diagram of PDL in-line monitoring and compensation. PC: polarization controller; OF: optical filter; PD: photodiode (narrow bandwidth)](image)

148
Figure 76 The root-mean-square (RMS) power variation of the monitored signal after passing through a 0.8-dB PDL component as a function of the polarization scrambling frequency.

PDL compensation is performed at each optical node between transmission fiber spans by adjusting the in-line compensators to minimize the monitored optical power fluctuation. Multi-channel compensation for PDL can be realized by scrambling the polarization state of either: (i) one of the data channel's itself, or (ii) a continuous-wave (CW), ancillary wavelength, thereby protecting the data channels from any unnecessary interaction that may occur between in-line PMD and polarization scrambling. As shown in Figure 77, if the PDL of a wavelength located at the center of the WDM channel range is completely compensated, then the adjacent channels will be partially compensated. Moreover, the broadband performance is determined by the PMD value of the optical module being
compensated, in which an increase in PMD will decrease the compensated bandwidth. If high PMD exists inside the individual devices of the PDL module, then the PDL compensator bandwidth will be reduced to perhaps single-channel operation (The number of channels depends on the channel spacing). Given ~0.2-ps of PMD in our PDL-generating in-line module, our compensator can correct for degradations over a 6-nm bandwidth. We note that the PMD value of the transmission fiber will not influence the PDL monitoring and compensation and, as a rough estimate, the PDL compensation bandwidth will be inversely proportional to the PMD that may occur between the PDL-generating device and the PDL compensator. Additionally, only one scrambler is required for all WDM channels at the transmitter side.

![Figure 77 Illustration of multiple-channel compensation as the center channel is completely compensated.](image)
4.6.2 System Demonstration Setup

Figure 78 (a) shows the experimental setup that includes a recirculating loop testbed that closely reproduces the Maxwellian statistics of PMD effects [41]. Four channels (1552 nm, 1554 nm, 1556 nm and 1558 nm) are NRZ modulated at 10 Gbit/s ($2^{15}$-1 PRBS) and decorrelated through 7 km of conventional single-mode-fiber (SMF). We chose to use an ancillary wavelength at 1555 nm for multi-channel PDL monitoring and compensation. The dispersion-managed recirculating loop consists of four EDFAs operating in the saturated regime, 82 km of SMF, and 12 km of DCF. The average PMD of 82-km SMF and 12-km DCF are 0.8 ps and 0.5 ps, respectively. In order to emulate the statistical distribution of PMD and PDL in real systems, the loop contains two loop-synchronized polarization controllers (PC), a polarization-maintaining (PM) fiber with ~5.4 ps DGD, and a PDL-generating element. The Jones matrix of each PC is updated after each round-trip interval of the loop to generate a series of random, uncorrelated polarization states. The PDL per loop, including the PDL-generating element, is ~0.85 dB. Figure 78(b) shows the measured instantaneous PDL distribution range as well as the average PDL value as a function of transmission distance. We determine from the data that the average PDL accumulates according to a square-root dependence. By tapping off some power of the signal and filtering out the desired wavelength, we can monitor the PDL value along the link and adjust a variable PDL compensating
module automatically. This tunable PDL compensating module is composed of two polarization controllers, with each one preceding a fixed PDL element. The loop testbed replicates an 800-km transmission link with \( \sim 2.4 \text{-dB root-mean-square PDL} \) and \( \sim 14 \text{-ps arithmetic average PMD} \) \((\text{RMS value} = \sqrt{8} \times 5.4 \text{ ps})\). We choose the 14-ps average PMD because this is the tolerance limit for the 10-Gb/s NRZ systems without PMD compensation according to the 30-min-per-year outage criterion [42]. The proposed technique applies also to the case with higher PMD value of the transmission fiber because PMD has no impact on the CW ancillary wavelength. The received optical SNR without PDL inside the loop is \( \sim 27 \text{ dB} \).
4.6.3 Compensation Results

Throughout the experiment, power penalties are measured by comparing the receiver sensitivity at a $10^{-9}$ bit-error-rate (BER) with the back-to-back one. Figure 79(a) shows the cumulative probability of power penalties for the 1556-nm channel with: (i) 14-ps average PMD (5.4 ps/loop) with no PDL, (ii) 2.4-dB RMS PDL (0.85 dB/loop) with no PMD, and (iii) both PMD and PDL combined as in cases (i) and (ii). In case (iii), the 2% tail of the penalty distribution dramatically increases from about 1.5 dB to 6.5 dB after introducing 2.4-dB PDL to the 14-ps PMD, and error floors occurred for 5 out of 500 samples. As a comparison, 43-ps DGD after
the same 800-km transmission with negligible PDL will induce ~4.0-dB power penalty [41].

![Diagram showing cumulative probability distributions of power penalty](image)

**Figure 79** (a) Cumulative probability distributions of the power penalty for three different cases: (i) 14-ps average PMD (5.4 ps/loop) with no PDL, (ii) 2.4-dB average PDL (0.85 dB/loop) with no PMD, and (iii) both PMD and PDL combined as in cases (i) and (ii). (b) Cumulative probability distributions of the power penalty after PDL compensation for two typical channels (1556 nm and 1558 nm). Each case in (a) and (b) obtained from 500 samples.

The resulting 500-sample power penalty probability distributions after in-line PDL compensation in the presence of 14-ps average PMD are shown in Figure 79(b). We use a scrambling frequency of 20 kHz. The 2% tail of the power penalty distribution is reduced to 1.4 dB and 1.9 dB for channels separated by 1-nm (1556 nm) and 3-nm (1558 nm) away from the ancillary wavelength, respectively. The residual penalties are induced mostly by the uncompensated accumulated PMD that takes a reasonably small value. A major wavelength-band limitation of PDL...
compensation is the PMD of the PDL module, which can induce wavelength-dependent PDL. In addition, compensation of wavelength-dependent PDL can be implemented on a per-channel basis by scrambling the SOP of each modulated data signal in the case of PDL components has high PMD values. We scramble the SOP of the 1556-nm data channel and reduce the 2% tail of the power penalty distribution to $\sim 2.0 \text{ dB}$. In the presence PMD, scrambling the data channel will incur higher penalties because the worst-case splitting between two PSP’s occurs more frequently.

In summary, we demonstrated in-line PDL compensation for optical modules that may exist high PDL. One may consider PDL compensation at the end of the link instead of performing periodical compensation. In this case, multiple-channel PDL compensation is difficult because the PDL compensation bandwidth is limited by the PMD of the whole transmission link. In addition, the interaction of many PMD and PDL elements along the link may necessitate a complicated compensator that accounts for the combined effects of PMD and PDL. A compensator using only PMD [45] or PDL elements may not be very effective to mitigate the channel degradation in this difficult case, which needs further investigation.
4.7 Programmable Group Delay Module Using Binary Polarization Switching

Delay generation is an important function that may be necessary for both optical and microwave communication systems. Optical delay lines, such as differential-group-delay (DGD) and true time delay (TTD) lines, are key elements in high performance transmission systems and networks.

As described before, DGD is characterized as the relative delay time between two orthogonal polarization states. In a high-speed optical transmission system, the non-ideal shape of the optical fiber core and mechanical stress on the fiber induce birefringence along the fiber and generate DGD, which in turn causes the annoying problem of polarization mode dispersion (PMD) [42, 128]. In general, the DGD and the principal axes of the fiber link depend on the wavelength and fluctuate in time as a result of temperature variations and external constraints [139]. Consequently, the corresponding pulse broadening caused by the PMD is random, both as a function of wavelength at a given time and as a function of time at a given wavelength. Unlike the effects of chromatic dispersion and fiber nonlinearity, which are deterministic and stable in time, the PMD-induced penalty can be totally absent at a given moment and adversely large several days later, causing an unacceptable bit-error-rate for no apparent reason. To ensure an acceptable outage probability for a fiber optic system, PMD compensation must be dynamic in nature.
and must be able to adapt to the random time variations. Therefore, the major challenges for combating the PMD effect include: (i) accurate and fast emulation of PMD statistics; and (ii) dynamic DGD generation for compensating the PMD produced in the fiber link.

Tunable DGD modules are highly desirable for both PMD emulation and compensation. As a PMD emulator, a tunable DGD module can provide a desired distribution mimicking the PMD fluctuation in a real fiber link within a short period of time and can be used to evaluate the performance of PMD compensators.

A typical first-order PMD compensator consists of a polarization controller followed by either a fixed or a variable DGD element. Variable DGD-based PMD compensators can reduce the risk of feedback loops being trapped in a locally optimized state and provide superior performance compared with fixed compensators [140]. In addition, in PMD compensators that use polarization scrambling at the transmitter to reduce the complexity and increase the stability of the feedback control, a variable DGD element is required in order to exactly cancel out the fiber’s DGD [52, 53, 131, 132]. As higher-order PMD effects become more significant, variable DGD elements may be required for higher-order PMD compensation as well [141-146].
The previous approach towards making a variable DGD element introduces a relative delay between two orthogonal polarization components after physically separating them with a polarization beam splitter. The two polarization components are then recombined with a polarization beam combiner [52, 147]. Such kind of device always has a low tuning speed (sub-second), large output polarization fluctuation, a large footprint, and poor control certainty due to mechanical motion.

In high-speed time-division-multiplexing (TDM) systems, a variable TTD (sometimes referred to as "group delay".) module can be used to precisely position data in an assigned time slot at the transmission end and select a desired time slot at the receiving end [148, 149]. In a microwave photonics system, a high-speed TTD module is desirable to perform microwave signal processing and phased array radar beam forming functions [150-152]. In these applications, a versatile TTD requires not only a large time delay (~microsecond) for wide dynamic range but also a fine delay resolution (sub-picosecond) to ensure accuracy. Because such a large time delay interval can be achieved with the optical path length switching approach, a reliable fine delay often becomes a challenge in device design. Several approaches [150-152], such as wavelength switching, have been proposed to achieve fine delay resolutions in a TTD. In 1994, X. Steve Yao et al first proposed photonic true time delay based on polarization switching [152] that can generate precisely controlled TTD and DGD, however, the concept was not demonstrated.
In this section, we demonstrate a compact, programmable photonic delay module based on the concept described in [152]. The module consists of six cascaded birefringent crystals separated by magneto-optic (MO) polarization switches. The lengths of the birefringent crystals are arranged in a binary power series, increasing by a factor of 2 in each section. The unique digital tuning mechanism ensures a precise and repeatable delay control to generate any DGD value from $-45$ ps to $+45$ ps in less than 1 ms with 1.40-ps resolution ($< 0.1$ ps resolution can be readily achieved by design). When used as a TTD, the module can generate any TTD value from 0 to 45 ps with a resolution of 0.68 ps.

### 4.7.1 Operating Principle

The tunable DGD module employs a polarization-switching approach to generate different delays [152]. As shown in Figure 80(a), the device is composed of multiple switch/delay sections. Each switch/delay section consists of a birefringent crystal to generate a fixed amount of delay and a MO polarization switch. The MO switch is internally made using a driving coil and garnet crystals. The lengths of the birefringent crystals are arranged in a binary power series, increasing by a factor of 2 in each section. Such a binary arrangement requires a minimum number of crystal sections and results in the highest possible delay resolution. The typical response curve illustrating the binary nature of the MO switch is shown in Figure 80(b). By driving the switch with current that is either less than the saturation current $I_1$ or
greater than saturation the current \( I_2 \), 90° polarization rotation/switching is achieved. Thus, at any switch/delay section, the input polarized beam can be either switched along the slow or fast axes of the birefringent crystal, corresponding to a longer or shorter delay. The total delay is the summation of the delays in each crystal segment and can be varied by the MO polarization switches.

\[ \delta \tau = \left| (n_e - n_o) \frac{\ell}{c} \right|, \quad (3.25) \]

where \( n_o \) and \( n_e \) are the ordinary and extraordinary indices of refraction, and \( c \) is the speed of light. Although the delay/switch sections can be placed in any order, we consider a case when the first bit (the least significant bit) in the switching

**Figure 80** (a) Illustration of a programmable DGD module based on polarization switching. (b) Typical response curve of a magnetooptic (MO) polarization switch.
command controls the shortest delay section and the last bit (the most significant bit) controls the longest delay section.

As mentioned in the introduction, the tunable delay module has two distinct applications: either providing variable DGD between two linear orthogonal polarization states or functioning as a photonic TTD by aligning the input SOP to one of the eigen polarization axes of the crystal set.

When the module is utilized to produce variable time delay to orthogonal polarizations, i.e., tunable DGD, the focus is on the time delay between the two eigen-polarization states. In this situation, when the n-th bit is switched from 0 to 1 state, the differential time delay of the n-th section changes its sign. The total differential time delay for an arbitrary binary state becomes

$$\Delta \tau_d = -\delta \tau \sum_{n=1}^{6} (-1)^{b_n} 2^{n-1}.$$  \hspace{1cm} (3.26)

where \(b_n(=0, 1)\) is the binary value of the n-th bit, determined by the polarization switch associated with the nth bit. The delay resolution is twice the unit delay time, i.e. \(2\delta \tau\), obtained by switching the least significant bit in Eq. 3.26. The differential group delay can be a negative number due to the sign flip of the delay time during switching. In a 6-bit module, a total of 64 DGD values (from -63\(\delta \tau\) to +63\(\delta \tau\)) can be generated with a resolution of \(2\delta \tau\).
When the variable delay module is used to provide TTD, the total absolute delay time $\Delta \tau_a$ at an arbitrary binary state is,

$$\Delta \tau_a = \delta \tau \sum_{n=1}^{6} \left[ 1 - (-1)^{b_n} \right] 2^{n-2} + \tau_0,$$

where $\tau_0$ is the constant bias time delay of the common optical path. Eq. 3.27 shows that when the n-th bit is switched from 0 (low) to 1 (high), the time delay of this section switches from the "off" state to the "on" state. In a 6-bit module, 64 different delay combinations from $\Delta \tau_a = \tau_0$ to $\Delta \tau_a = 63\delta \tau + \tau_0$ can be generated with a delay resolution of $\delta \tau$ (as defined in Eq. 3.25).

Based on the principle described in Figure 80, we designed and fabricated a novel programmable delay module using 6 birefringent crystals (i.e. a 6-bit module) that can generate tunable DGD values from $-45$ ps to $+45$ ps with a resolution $(2\delta \tau)$ of 1.40-ps, or tunable TTD values from 0 to 45 ps with a resolution $(\delta \tau)$ of 0.7-ps. One of the most promising features of this module is the sub-millisecond switching speed desirable for PMD compensation, PMD emulation, as well as microwave signal processing applications.
4.7.2 Static Characteristics

4.7.2.1 DGD Characterization

We carefully measured the static characteristics of the tunable delay module for DGD generation. To verify the accuracy of generated DGD values of the module, we used a commercial PMD analyzer to measure both the DGD value and the second-order PMD at each logical DGD state. As shown in Figure 81, the measured DGD values agree well with the designed DGD values and each DGD value is exactly reproducible. Equally as important, the second-order PMD is very small, less than 85 ps$^2$, even at high values of DGD.

![Figure 81. Measured DGD and the second-order PMD as a function of designed DGD.](image-url)
Insertion loss, wavelength-dependent loss and polarization-dependent loss (PDL) are key figures of merit of any in-line optical component. Figure 82(a) shows the measured insertion loss as a function of wavelength across the C-band (1535~1560 nm) using a broadband amplified spontaneous emission (ASE) source, and Figure 82(b) shows the measured PDL values for all designed DGD states. The insertion loss is \( \sim 1.3 \) dB with a wavelength dependent variation of \( < 0.15 \) dB. The PDL values range from 0.02 dB to \( \sim 0.28 \) dB for all the DGD states.

![Figure 82](a) Insertion loss of the delay module as a function of wavelength (1535~1560 nm) at a typical DGD state (~ 40 ps). (b) Polarization dependent loss (PDL) as a function of designed DGD values.

### 4.7.2.2 TTD measurement

When our tunable DGD module is utilized as a photonic true time delay line, i.e. providing absolute delay, the minimum delay occurs when all \( b_n \) values are equal to 0 (000000 state) in Eq. 3.27, and the maximum delay occurs when all \( b_n \) values are equal to 1 (111111 state). The total time delay and the delay resolution are proportional to the unit delay, and one can select appropriate birefringent crystals
to achieve delay resolutions down to the femtosecond level for ideal fine delay tuning.

Figure 83 True time delay measurement setup using an ASE broadband source. OSA-optical spectrum analyzer.

The absolute time delay or TTD is measured experimentally using a fiber-optic Mach-Zehnder interferometer, as illustrated in Figure 83. Using a broadband source centered at 1550 nm as the input light, the interferometer output intensity can be expressed as

$$I_o = I_i \left[ 1 + m \cos \left( 2\pi f (\Delta\tau_a - \tau_1) \right) \right]$$

where $f$ is the optical frequency, $m$ is the optical modulation index, $\Delta\tau_a$ is the absolute delay time defined in Eq. 3.27, and $\tau_1$ is the absolute delay time in interferometer arm 1. The interferometer arm 1 is first adjusted so that the absolute delay in arm 1 is equal to the common path delay $\tau_0$ of the tunable delay module. The adjustable part of the absolute time delay is then measured by counting the number of fringes over the specified optical frequency range at different binary
switching states. Assuming there are \( M \) fringes over a bandwidth of \( \Delta f \), the delay can be calculated as

\[
\Delta \tau_a = \frac{M}{\Delta f}
\]  

(3.29)

To improve measurement accuracy, we use an adjustable delay line to maintain the fringe number \( M=0 \) on the optical spectrum analyzer. The absolute time delay was derived from the displacement change of the variable optical delay line. The measured data agrees well with the theoretical prediction, as shown in Figure 84.

![Figure 84](image)

**Figure 84** Comparison between the measured absolute true time delay and designed values.

### 4.7.3 First-order PMD Emulation

The precise and repeatable DGD generation capability of our DGD module is ideal for generating a series of DGD values with any statistical distribution (e.g. Maxwellian, Gaussian, or Lorentz) for a given number of samples. We therefore developed the software tools necessary to control the DGD module so it can be
used to generate statistical DGD samples with a Maxwellian distribution and a selectable average DGD value $\Delta \tau$. The tunable average DGD value for the Maxwellian distribution can be determined by using [153].

$$\sqrt{\frac{2}{\pi}} \frac{\Delta \tau^2}{\alpha^3} \exp \left[ -\frac{\Delta \tau^2}{2\alpha^2} \right]$$

where $\alpha = \frac{\langle \Delta \tau \rangle}{\sqrt{8 / \pi}}$.  

Figure 85(a) shows the measured distribution for 500 samples with an average value of 10 ps, and Figure 85(b) shows the instantaneous DGD of the samples. Although this is only first-order PMD emulation (almost no higher-order effects involved), it has a much higher speed than standard PMD emulator and can be precisely controlled. In addition, its features enable users to emulate pure or statistical all-order PMD distributions [154-156]. These characteristics should prove powerful in evaluating the dynamic performance of PMD compensators.

Figure 85 Repeatable DGD generation is used to generate dynamic first-order PMD distribution with tunable average values. (a) First-order PMD distribution of 500 samples with an average DGD of 10 ps. (b) Instantaneous DGD values of the samples
4.7.4 Dynamic Performance

In continuous data traffic network applications, it is critical that the system performance be unaffected during DGD state switching. Therefore, the dynamic performance of this DGD module must be well understood and controlled.

In our experiment, we first characterize the effect of switching on the DGD values. In principle, when the device is switched from one DGD state to another, the DGD value will change from one value to another precisely. However, because the device has a finite switching speed, the DGD value during switching is different from either the initial or final states. When the device is switched in smallest step ($2\delta\tau$), the maximum DGD value excursion from the DGD value of the ending state is defined as the transient DGD. As it is difficult to directly measure the transient DGD during switching due to the limited response speed of the measurement instrument, we used a quasi-static measurement method. Instead of having a full 90° polarization rotation in one step to change the DGD value from one state to another, we incrementally change polarization rotation angles in steps of a few degrees (as shown in Figure 86(a) by increasing or decreasing the control current on the MO switches) while measuring the corresponding DGD with a commercial polarization analyzer. We repeat the procedure for all DGD values from 0 ps to 45 ps in steps of 1.40 ps with the results shown in Figure 86(b). The insert shows the detailed DGD values between two adjacent DGD states. It is evident that the
transient DGD is always less than the step size. The small transient DGD is critical when the device is used in PMD compensators. Note that there will also be 2nd-order PMD involved in during the switching. The value of 2nd-order PMD can be up to several-hundred ps² depending on the length (or DGD) of the crystals.

![Diagram](image)

**Figure 86** Characterization of transient DGD effects during polarization switching. (a) Quasi-static measurement of transient DGD by step-varying the control current of the MO switches (b) Measurement of transient DGD by equivalent small polarization variation between different DGD values from 0 to 45ps.

Another transient effect is the loss variation during fast polarization switching. Through proper mechanical alignment and electronic circuit design (e.g. proper arrangement of the switching sequence), this transient loss can be reduced to < 0.6 dB, even in the case when all six bits are switched at the same time, while maintaining a total switching speed of less than 1 millisecond.
The most straightforward and effective evaluation of a fiber optic device is to test the device in an actual fiber optic link. In order to test the impact on a system due to the transient effects, we integrate the DGD element into a 10-Gb/s NRZ transmission link modulated at $2^{31}$-1 PRBS, as shown in Figure 87(a). The input optical-signal-to-noise-ratio (OSNR) to the delay module is set to 30 dB (0.1 nm bandwidth) by changing the input power into the first erbium doped fiber amplifier (EDFA). An optical pre-amplifier before the receiver is used to increase receiving sensitivity. The system’s back-to-back sensitivity is measured to be -31 dBm. Power penalties are measured by comparing the receiver sensitivity of the system at a $10^{-9}$ bit-error-rate (BER) with the back-to-back sensitivity. We measured two cases: (i) power penalties when the delay module is set at different static DGD values, and (ii) power penalties when neighboring states of the module are jogging back and forth at 1 kHz (1 ms continuous switching). As shown in Figure 87(b), a negligible power penalty of $< 0.2$ dB due to the fast polarization switching (jogging) was obtained. This result assures that the transient effects of the delay module will be too small to cause any concerns in a real system and that the device can be effectively used in systems for PMD compensation, emulation, or signal processing applications without causing harmful side effects. We also measured the transient effect for the case in which there is no optical pre-amplifier before the receiver and found that the transient-induced power penalty is less than 0.5 dB when jogging at 1 kHz. Such a power penalty is mainly due to transient loss during switching.
Figure 87 System evaluation of the tunable DGD module. (a) Experimental setup of a 10-Gb/s NRZ ($2^{31}-1$ PRBS) link. (b) Measured power penalties for two cases. Open circle: the DGD module remains static after changing state; solid circle: the DGD module keeps jogging back and forth between neighboring states at 1 kHz.

4.7.5 Applications

As mentioned in the introduction, the dynamic delay module presented in this paper has numerous applications in both fiber optic communication systems and microwave systems.
4.7.5.1 **High-speed PMD emulation**

PMD emulation is important for evaluating systems’ PMD susceptibility and for developing PMD compensators. A fast and repeatable DGD generator is particularly useful for testing the response time and effectiveness of a PMD compensator under different PMD varying conditions.

There are several approaches to making PMD emulators, such as those based on splitting the input light into two polarization components and delaying one of the components by free-space optics, those based on rotating the relative angles of several sections of birefringence crystals [157], and those based on heating or rotating different sections of PM fiber [40, 158]. All of these PMD emulators suffer from a slow response time (on the order of seconds) and poor repeatability, making them less effective in evaluating PMD compensators. In contrast, because of its high speed and precise DGD repeatability, the delay module reported in this paper is ideal for such applications. For example, our DGD module is capable of generating repeatable DGD variation as fast as 200 microseconds and thus can be used for evaluating the impulse response of a PMD compensator that has a speed on the order of one millisecond.
4.7.5.2 First and higher-order PMD compensation

As previously mentioned, a tunable DGD module is a key component in many PMD compensation configurations for first and higher-order PMD [141-146]. In a recent demonstration of PMD compensation scheme, a variable DGD element is placed at the receiver to completely compensate the first-order PMD [142]. Some typical first and higher-order PMD compensation schemes using a variable DGD module are shown in Figure 88.

Figure 88 Different approaches for PMD compensation using variable DGD modules.
As a tunable DGD module based on polarization splitting and combining suffers from slow tuning speeds as well as a large footprint and vibration sensitivity, it is not suitable for field deployment. Finally, the output polarization stability of such a device is poor due to length fluctuations of the independent optical paths of the two polarization components. Such polarization instability complicates the PMD compensator feedback loop design and may cause the loop to be unstable.

On the other hand, the delay module reported in this paper has the advantage of compact size and high speed, making it field deployment friendly. In addition, its output polarization is much more stable than that of the DGD module described above because here the two polarization components share the same optical path. Finally, the binary DGD tuning mechanism ensures that the module can be tuned to any desired DGD value precisely and quickly without any feedback control, eliminating the need of another DGD-control loop.

4.7.5.3 Multi-wavelength laser source with tunable channel spacing

Multi-wavelength lasers have received considerable attention recently due to their potential application to optical measurement, optical sensors, and wavelength-division-multiplexed systems. A PM fiber placed inside a ring laser cavity can function as a multi-wavelength selector to produce a multi-channel optical output [159, 160]. However, such line spacing is fixed once the length of the PM fiber is
selected. To make a multiwavelength laser with tunable channel spacing, a tunable DGD module can be used. In principle, the channel spacing $\Delta f$ is related to DGD value of the module by $\Delta f$ (Hz) = 1/DGD (ps). Experimental verification of this method using our variable DGD module is shown in Figure 89 [161].

![Figure 89 Output channel spacing as a function of DGD for a multiwavelength fiber ring laser.](image)

4.7.5.4 Bit alignment in TDM systems

In high-speed time-division-multiplexed (TDM) systems, it is critical to precisely position data in an assigned time slot at the transmission end and to select a desired time slot at the receiving end [148, 149]. As shown in Figure 90, the delay module reported here can accomplish this function easily. Data from transmitter 1 (TX1) and transmitter 2 (TX2) are combined by a polarization combiner (PBC) before entering the variable DGD module. The signal from TX1 is aligned with one eigen-polarization axis of the device and the signal from TX2 is aligned with the
other. The relative group delay between the two signals can be delayed precisely and thus can align the data bits from the two transmitters in time. At the receiver side, the two channels can be demultiplexed via a polarization beam splitter (PBS) placed after a polarization stabilizer.

![Figure 90 Illustration of data bit alignment using the delay module.](image)

### 4.7.5.5 Beam steering of phased array radar systems

For airborne and space-based phased array radar systems operating at mm-wave frequencies (20 GHz and above), fast beam steering with a fine scan resolution is required [162]. One effective method of steering the radar beam is to feed each radiation element on the radar panel with optical fibers and change the relative time delay of the signals feeding to the radiation elements, as shown in Figure 91. As mentioned earlier, the delay module reported in this paper can be used as a TTD for steering the phased array radar [163]. This programmable TTD can be readily modified into a two dimensional device and has the advantages of high packing
density, low loss, and easy fabrication. The delay resolution of the device is sufficiently fine for accurate beam steering, and the total delay can be made adequately large by increasing the length of each crystal section to cover the desired scanning angles. This device can also be simplified to a phase-shifter/beam-former for phased arrays of narrow bandwidth where true time delay is not necessary.

Figure 91 Illustration of using the variable TTD module for steering a phased array radar. The relative delays between the radiation elements are adjusted by the variable TTD to cause the total radiation wave front to tilt and hence effectively scan the angle of the radar beam.
4.7.5.6 Transversal filter

Fast and widely tunable filters are important for signal extraction and processing in spread spectrum microwave communication and radar systems [164-166]. Transversal filters based on photonics technology have several attractive features for such applications, including wide bandwidth, high packing density, low loss, low weight, remote capability, and immunity to electromagnetic interference. True time delay is a critical element in a transversal filter, as shown in Figure 92. The fast switching speed, high delay precision, and high delay repeatability of the delay module reported here are among the most attractive features for a transversal filter. These features in turn make the filter high speed, precise, and highly repeatable.

![Figure 92 Illustration of using the variable TTD module in a transversal filter. The output signal is the coherent summation of all branches. The relative phase (delay time) and amplitude changes in the branches alter the shape of the output signal.](image)

In summary, due to the unique binary polarization switching design, the delay module also exhibits both excellent static and dynamic characteristics, as
summarized in Table 3. With a special emphasis on its dynamic figures of merit for polarization mode dispersion (PMD) emulation and compensation applications, we demonstrated that the device exhibits a negligible transient-effect induced power penalty (<0.2 dB) in a 10-Gb/s NRZ system. It can be readily modified for a 40Gb/s network by reducing the lengths of the birefringent crystals.

Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Typical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insertion loss</td>
<td>1.3±0.2 dB</td>
</tr>
<tr>
<td>PDL</td>
<td>&lt; 0.28 dB</td>
</tr>
<tr>
<td>Switching time</td>
<td>&lt; 1 ms</td>
</tr>
<tr>
<td>Optical return loss</td>
<td>&gt;50 dB</td>
</tr>
<tr>
<td>TTD range</td>
<td>0~45 ps</td>
</tr>
<tr>
<td>DGD range</td>
<td>-45~45 ps</td>
</tr>
<tr>
<td>TTD resolution</td>
<td>0.7 ps</td>
</tr>
<tr>
<td>DGD resolution</td>
<td>1.40 ps</td>
</tr>
<tr>
<td>Transient loss</td>
<td>&lt; 0.6 dB</td>
</tr>
<tr>
<td>Transient DGD</td>
<td>&lt; 1.40 ps</td>
</tr>
<tr>
<td>WDL</td>
<td>&lt; 0.15 dB</td>
</tr>
<tr>
<td>Switch voltage</td>
<td>TTL (5V)</td>
</tr>
</tbody>
</table>
4.8 Polarization-mode-dispersion emulator using variable
differential-group-delay (DGD) elements and its use
for experimental importance sampling

To accurately characterize the outage probability of networks that may or may not incorporate PMD compensation, it is essential to have a PMD emulator that can quickly cycle through the various PMD states expected in an optical fiber.

Previously demonstrated PMD emulators are typically constructed using several randomly-coupled polarization maintaining (PM) fibers [40][158] or birefringent crystals mounted on rotation stages [157]. Two major drawbacks of current emulators are: (i) the lack of stability or repeatability, and (ii) the inability to vary the PMD statistics (i.e., no tunable average DGD). In general, emulator repeatability is limited by the environmental sensitivity of the birefringent elements and/or the poor control certainty of any mechanical parts. Moreover, the average DGD of these emulators is fixed and cannot be reconfigured to emulate different fiber plants.

System designers typically require that system outages (penalty >1dB) due to PMD occur with a probability of $10^{-6}$ or less (<1 min/yr) [128]. To assess the effects of PMD on a system, both with and without compensation, PMD emulators are used to cycle through different PMD states. However, it is very difficult to characterize
system outage probabilities using previously reported PMD emulators, or even with computer simulations, because of the extremely large number of randomly generated PMD states that must be explored to obtain a reliable estimate.

Importance sampling (IS) is a powerful tool for obtaining very low probability events with relatively few sample points [167]. This is accomplished by altering the method of obtaining the random samples to concentrate the measured results in the area of interest in the sample space. This will distort the probability distribution of the measured results, so each sample must then be appropriately weighted to map the measured values back onto the proper distribution function.

Thus far, importance-sampling techniques for PMD emulation have only been accomplished using computer simulations [168-171]. This is because a critical drawback of most previously reported PMD emulators is that they do not possess the programmability, or stability, required to perform IS. To perform IS with these emulators requires deterministic control of the coupling angle between the PMD vectors of adjacent sections in order to preferentially align them to obtain rare PMD events. This is extremely difficult to accomplish because the environmental sensitivity of the birefringent elements causes the direction of the PMD vectors to drift over time (even if the DGD remains constant, tiny variations in the birefringence will cause large changes in the PMD vector’s direction). Furthermore, even with highly stable elements, it would still be a significant
challenge to determine the PMD vector between sections and accurately produce the desired coupling angles for each sample. One recent publication shows another PMD emulation approach that may be a good candidate for such applications, though it has not yet been demonstrated for importance sampling [172].

In this section, using three programmable DGD elements as described in the previous section, we experimentally demonstrate a high-speed (<1 ms), stable and repeatable PMD emulator that can generate any desired Maxwellian DGD distribution, with an average up to 35 ps, and corresponding 2nd-order statistics. The stability and repeatability of the emulator DGD and output state of polarization (SOP) are characterized. Our emulator maintains a given PMD state over several hours, whereas the output SOP of other emulators drifts dramatically within minutes. A PMD variation of <5% is obtained for 50 samples repeated 4 times.

Using this emulator, we present a new method to readily enable experimental importance sampling to produce low-probability events without the need to determine and control the direction of the PMD vector between sections. With this emulator, we show that importance sampling can be accomplished by simply biasing the distribution of DGD values applied to each element, as opposed to controlling the coupling angles between sections. As such, only uniform scattering of the polarization coupling between sections is required, which is easily
accomplished with electrically driven polarization controllers. Here we experimentally use importance sampling to efficiently obtain rare, Maxwellian distributed DGD events with probabilities as low as $10^{-24}$ (for $\langle \text{DGD} \rangle = 15$ ps) and correspondingly rare 2\textsuperscript{nd}-order PMD events after taking only 1,000 samples. We also employ “multiple importance-sampling” techniques to combine the results from three different distributions to achieve better coverage of the entire sample space. The resulting distribution tail extends to $10^{-30}$.

In addition, experimental measurements have been performed in the past to characterize the system Q degradation due to PMD [173]. However, an impractically large number of random samples must be taken to explore the rare events in the distribution tail using these previous methods. Therefore, we also use multiple importance-sampling to measure the Q degradation due to the PMD generated by our emulator. The measured Q-penalty probability distribution extends to <$10^{-17}$ with only 1800 experimental samples.
4.8.1 PMD Emulation with Tunable Statistics

As shown in Figure 93, the emulator is constructed from three variable DGD elements separated by two fiber-squeezer-based polarization controllers. Several variable DGD generation approaches have been proposed [174, 175], and here we employ a very practical approach that was described recently in [176]. Each variable DGD element consists of several birefringent crystals whose lengths increase in a binary series and are separated by electrically driven polarization switches. The elements can be digitally programmed to generate any DGD value from -45 ps to +45 ps with a tuning speed of <1 ms and a resolution of 1.40 ps. This resolution is a consequence of the structure of the DGD of sections included in each variable DGD element [176]. A computer is used to control the emulator to
randomly generate any desired DGD distribution for each element and to uniformly scatter the polarization between sections [177]. To obtain a Maxwellian DGD distribution at the emulator output, the DGD values of each element are varied according to a Maxwellian distribution with average, $\Delta \tau$ [178][179]. This yields an average DGD of $3^{1/2}(\Delta \tau)$ for the total emulator and an average 2nd-order PMD distribution that has the correct shape but falls slightly short of that expected for a real fiber, as shown in a recent simulation result [179].

To demonstrate tunability of the PMD statistics, three different distributions are generated, as shown in Figs. 94 and 95 for $<\text{DGD}> = 10$, 25 and 35 ps. As expected, the DGD values closely match the expected Maxwellian distribution. The corresponding 2nd-order PMD distributions have averages of 31, 174, and 322 ps$^2$, which are ~30% lower than expected for a real fiber, and also lower than the expected values in the recent simulations [179]. All of the PMD measurements shown throughout this paper were performed using the Jones matrix method on a commercial PMD analyzer [180].
Figure 94 Three output DGD distributions with different statistical averages, each showing a good fit to the Maxwellian pdf (probability density function) expected from a real fiber, (a) linear scale and (b) logarithmic scale.

Figure 95 The corresponding 2\textsuperscript{nd}-order PMD distributions to Figure 94, (a) linear scale and (b) logarithmic scale. The solid curves are drawn from computer simulation results of this emulator, showing that the mean 2\textsuperscript{nd}-order PMD is about 30\% lower than that expected from a real fiber with the same average DGD since only 3 sections are used.

4.8.2 Emulator Stability and Repeatability

Stability and repeatability are highly desirable features for PMD emulators as they enable one to examine system performance at specific PMD conditions and to
achieve deterministic control of the emulator's state. To characterize stability, we observed the output SOP variation of our emulator in a laboratory environment. SOP stability is important because it indicates that the direction of the PMD vector remains stable, which is a necessary condition for repeatability. Figure 96(a) shows that the output SOP of our 3-section emulator remains nearly constant over a 4-hour period. For each individual section, we observed that the SOP varied negligibly over tens of hours. In order to characterize the stability of our emulator, we measured the variation of the emulator’s DGD and 2\textsuperscript{nd}-order PMD over 30 minutes for both high and low DGD values. As shown in Figs. 96(b) and (c), the DGD remained remarkably stable for both cases, and the 2\textsuperscript{nd}-order PMD varied within a reasonable range (e.g., for the sample with 72-ps DGD, the DGD varied <5% and the 2\textsuperscript{nd}-order PMD varied ~15% over 30 minutes). It should be noted that some of this variation is due to the inherent measurement error of the PMD analyser system (the same measurement for a single piece of PM fiber with 50 ps DGD also showed ~5% variation over 30 minutes).
Figure 96. (a) Output SOP stability of the emulator over 4 hours. (b) and (c) measurements of the DGD and 2\textsuperscript{nd}-order PMD over 30 minutes for two different average DGD values: (b) $<\text{DGD}> = 71$ ps, (c) $<\text{DGD}> = 5$ ps.

To characterize the SOP repeatability, the emulator was repeatedly cycled six times through five different DGD states at 1-minute intervals. Figure 97(a) shows that the output SOP repeatedly returns to the same point on the Poincaré sphere for each DGD state (The traces shown in the figure may not illustrate the real phase or PMD vector variations during polarization switching due to the relatively slow response of the polarization analyzer). To characterize the DGD repeatability, the emulator was cycled through 50 different sets of control parameters four times. The total test time was $\sim 1$ hour. The 50 measured DGDs from the four tests are overlaid in Figure 97(b). At each sample point, the DGD variation is typically $<5\%$, indicating the ability to generate a look-up table of control parameters and corresponding DGD output values.
Figure 97 (a) Repeatability of the output SOP as the emulator is repeatedly tuned to 5 DGD states. (b) Four repeated measurements of 50 DGD samples are overlaid to show the repeatability of the output DGD (variation is typically <5% at each sample point with some of this variation due to measurement error).

Although the above results indicate that this emulator is a highly promising candidate for applications that require the PMD state to remain stable over several hours, it is still possible for the output state to vary with large temperature changes (more than a few degrees), or if the single-mode fiber pigtails between the sections are perturbed. Therefore, a simple and effective method to achieve repeatability without concern about emulator drift would be highly desirable. To accomplish this, three in-line polarimeters are inserted after each section and a polarization controller is added after section three (Figure 98). The polarimeters are used to record the SOP between sections for different emulator states. Since the DGD of each section is known and extremely stable (<0.1ps/80°C), the additional SOP information allows us to construct a lookup table of output first and second-order PMD vectors versus the six input DGD and SOP parameters. After recording the
input and corresponding output parameters for each randomly-generated sample during a long system test, the operator can return to any previously recorded PMD state (e.g. one that caused high penalty) for further investigation by simply adjusting the DGD elements and polarization controllers to re-acquire the set of input values for that sample. Even after environmental or polarization-coupling perturbations, this table can always be used to re-acquire a desired PMD state using automatic feedback control of the polarization controllers to obtain the needed SOP coupling.

Figure 98 Emulator setup incorporating simple, in-line polarimeters with automatic feedback to control the SOP between sections. This enables the generation of a lookup table of input DGD and SOP values versus the corresponding output 1st ($\tau$) and 2nd-order PMD ($\tau_\omega$).

The effectiveness of this concept is shown in Figure 99(a). The input parameters were recorded for five “original” PMD states. The fiber pigtails were then perturbed between sections, causing the SOPs, and therefore, the output values, to change. The polarization controllers were then adjusted to re-acquire the previously recorded SOPs for each sample, causing the output DGD and 2nd-order PMD values to return to their original states.
This setup may also be used to easily control the PMD vectors of the three sections to obtain any desired PMD state. For example, in order to generate pure first-order PMD, a one-point calibration is performed to measure the three SOPs that maximize the output DGD for a set of three input DGD values. As the DGD elements are varied, the SOPs can be controlled back to this calibration point to re-align the vectors and obtain the maximum output DGD (and negligible $2^{\text{nd}}$-order PMD) as shown in Figure 99(b). This same concept is used to set fixed $90^\circ$ angles between the PMD vectors of each section to generate large values of $2^{\text{nd}}$-order PMD (Figure 99(c)).

This approach of using polarimeters to control the polarization coupling between sections could be useful for applications requiring deterministic control of the PMD vectors between sections. These include importance sampling [157, 167] and the generation of accurate higher-order statistics with only three sections by dialing-in the desired first and $2^{\text{nd}}$-order PMD vectors, instead of randomly generating these samples [168, 169].
**4.8.3 Importance Sampling using Programmable DGD Elements**

Importance sampling (IS) is a well known technique for biasing the method of obtaining random samples such that the statistical results are concentrated in an area of interest in the sample space. This allows one to more effectively study the effects of a random phenomena, such as PMD, with fewer trials than would ordinarily be required by using conventional Monte Carlo techniques. Using the emulator described in the previous section, we are able to apply this powerful technique to physical fiber systems so that the impairments due to rare PMD events can be quickly and experimentally characterized and provide a comparison for results obtained previously via computer simulations.
The importance sampling technique we employed is conceptually illustrated in Figure 100. We exploit the programmability of the DGD elements to perform IS by applying randomly selected DGD values from a probability density function (pdf) other than a Maxwellian. Any pdf may be used, but the best choices are those that will tend to generate more output samples in the region of interest with the fewest possible measurements. For our first case, we chose to apply a uniform distribution of DGD values to each element over their full 45-ps range. In contrast to conventional importance sampling techniques, deterministic polarization coupling (i.e. biased polarization coupling) between sections is not required in this new approach. Here we still only apply uniform polarization coupling between sections.
The DGD applied to each element and the corresponding output DGD and 2\textsuperscript{nd}-
order PMD are recorded for each sample. As expected, the measured output values
will not follow the desired Maxwellian distribution and must be properly weighted
to adjust their probabilities to match the desired Maxwellian statistics. For each
DGD section, let $p(x_i)$ be the probability of obtaining DGD $x_i$ using the desired
Maxwellian pdf (with an average DGD of $\Delta\tau = <\text{DGD}>/(3^{1/2})$) and $p^*(x_i)$ be the
probability using the uniform pdf. For each sample, $i$, three likelihood ratios,
$p(x_i)/p^*(x_i)$, are computed using the three applied DGD values for the $x_i$s. The
three ratios are multiplied together and divided by the total number of samples to
determine the “weight” for each sample. The output DGD values are then sorted,
while keeping track of the corresponding weights. The DGDs and corresponding
weights are grouped into DGD bins and the weights in each bin are summed to
obtain the probability for that bin. These probabilities are then plotted alongside a
Maxwellian, integrated over each bin, for comparison. Note that, since the
programmable ability plays a key role for importance sampling using biased
distributions, the stability and repeatability of programmable DGD elements, which
are highly desirable in conventional PMD emulators to facilitate long-term system
evaluation, are not crucial for this application.
4.8.4 Experimental Importance Sampling Results

Figure 101 Importance sampling results for 1000 uniformly distributed DGDs applied to each section (0 to 45 ps). (a) Measured output pdf (note several values at large DGDs are generated), (b) pdf after renormalizing the data to obtain a Maxwellian distribution with $<\text{DGD}> = 15$ ps. Points down to $10^{-24}$ are generated with only 1000 samples. The inset shows the linear scale.

Figure 102 2nd-order PMD pdf for the importance sampling experiment described in Figure 101 (a) Measured output pdf and (b) pdf after renormalizing the data. The 3-section emulator produces a 2nd-order pdf with the correct shape, but a slightly lower average than that of a real fiber because of the small number of sections. The inset shows the linear scale.
The resulting DGD and 2\textsuperscript{nd}-order PMD probability distributions when 1000 uniformly distributed DGDs are applied to the three sections are shown in Figs. 101 and 102. Figs. 101(a) and 102(a) show the distributions of the unprocessed, measured values. Clearly, numerous large DGD and 2\textsuperscript{nd}-order PMD values result, relative to the unbiased case. In Figs. 101(b) and 102(b), the measured samples have been renormalized as described above, where $p(x_i)$ is a Maxwellian distribution with $\Delta \tau = 8.7$ ps/section. As expected, the experimental points for the total DGD closely approximate a Maxwellian with $<\text{DGD}> = 3^{1/2}(8.7) = 15$ ps and rare events down to $10^{-24}$ are obtained, whereas conventional sampling would only reach $10^{-3}$ probabilities with 1000 trials. The experimental 2\textsuperscript{nd}-order PMD pdf (Figure 102(b)) has the correct shape, but falls short of the theoretical pdf for a real fiber because only three sections are used. However, it is notable that large 2\textsuperscript{nd}-order PMD values are obtained with this method.

To efficiently obtain PMD events covering the entire range from low to high values, we used the technique of “multiple importance sampling” [169] to combine the results of several experiments using different DGD pdfs applied to each section.
Figure 103  Multiple importance sampling results. (a) The three DGD distributions applied to each section (840 samples/distribution). (b) Resulting DGD distribution showing that each pdf generates samples in different regions to cover the entire Maxwellian. (c) Resulting 2nd-order PMD pdf. The insets show the pdfs on a linear scale where it is evident that better coverage of lower values is achieved in comparison to the case shown in Figs. 101 and 102.

As shown in Figure 103 (a), an unbiased, Maxwellian pdf was used to obtain several values in the low-DGD region, a negatively sloped linear pdf was used to obtain low to medium DGDs, and a positively sloped pdf was used to obtain high DGDs. 840 samples were taken for each distribution. The experimental results are weighted as described in [169] to obtain the distributions shown in Figs. 103(b) and
(c). The multiple IS technique provides better coverage of the entire sample space. The resulting distribution tail extends to $10^{-30}$.

### 4.8.5 Measurement of Q degradation

In addition to characterizing the PMD statistics of the emulator, it was inserted into a 10-Gb/s transmission system and the multiple IS technique was used to characterize the impact of the PMD on the statistics of the system Q for both average and extremely rare PMD events. The Q values are measured at the optically pre-amplified receiver for 10-Gb/s NRZ data using a $2^{23}-1$ PRBS. The Q measurement is performed using the method presented in [181].

Similar to the method used in the previous section, importance sampling for Q measurements is accomplished by applying a biased DGD distribution to each emulator section (chosen to emphasize the region of interest) and then appropriately weighting the Q penalty results to obtain the proper probability density function (pdf). Instead of applying Maxwellian-distributed DGDs to each section (the conventional, unbiased case), a uniform distribution is applied to each section to cause the emulator to generate more samples at high DGD values, which often correspond to low Q values. The applied DGDs and measured Q values are recorded for each sample point and the measured Q probabilities are then appropriately corrected for the effects of biasing the DGD distributions.
Figure 104 (a) The DGD distributions applied to each section for the measurements of Q degradation: unbiased (Maxwellian, with average DGD ~ 8.7 ps per section) and biased (uniform). (b) Measured Q probabilities using multiple IS for a system with average DGD = 15 ps (~10^{-17} prob. that Q drops to ~12 dB).

To achieve good coverage of both low and high probability events, we also employed multiple IS here, in which the results from two IS experiments were combined as shown in Figure 104(a) for an unbiased Maxwellian (~750 samples) applied to each section and a biased uniform distribution (~1000 samples). The unbiased distribution provides good coverage of high probability (average) events, whereas the biased case yields a large number of low-probability values since a large number of high-DGD samples are generated. The resulting Q probability distribution is shown in Figure 104(b) for a system with 15-ps average DGD (linear scale is shown on inset). The Q degraded to ~12 dB at a probability of ~10^{-17}, and a complete loss of the signal was observed in several samples (not plotted here). This is an extremely low-probability event that occurs when the PMD is ≥1 bit time. It should also be noted that these Q measurement include the effects of both the first
and higher-order PMD, although the mean of the 2\textsuperscript{nd}-order PMD distribution of this emulator is ~ 30% lower than that of a real fiber with the same mean DGD because only three tunable DGD elements are used.

To summarize this section, an electronically controllable PMD emulator that is constructed from three programmable DGD elements has been experimentally characterized and used to study the effects of PMD on a fiber transmission system. The generated PMD statistics of the emulator can be readily tuned by simply applying different DGD distributions to each section. The stable and repeatable DGD programmability of the emulator enables the experimental realization of importance sampling, a powerful technique that allows system designers to investigate extremely low probability events that may cause system outages for only minutes per year with relatively few random samples. The PMD emulator's statistics, performance in terms of stability and repeatability, and use for a system characterization in terms of PMD-induced Q degradation were evaluated. One of the limitations of this emulator is that it would require more DGD elements in order to accurately emulate the statistics of 2\textsuperscript{nd}-order PMD. However, we would also note that, as more sections are added, the efficiency of this importance sampling method (biasing the DGD distribution of each section toward higher values while uniformly scattering the polarization between sections) is likely to decrease because we expect that it will become less probable for the PMD vectors of the DGD elements to be aligned as the number of sections increases due to the uniform
polarization coupling. This is in contrast to the method commonly used in computer simulations where the PMD vectors of the sections are preferentially aligned while their DGDs remain fixed. While the technique presented here is experimentally simpler to employ and works well for first-order PMD, it is not yet clear what the tradeoff is between the importance sampling efficiency and the number of tunable DGD sections used in the emulator.

4.9 Differential Group Delay Monitoring Used as Feed Forward Information for Polarization Mode Dispersion Compensation

One of the critical challenges for next-generation high bit-rate optical fiber transmission systems (≥10 Gbit/s/channel) is polarization mode dispersion (PMD) [42]. As the birefringence of a fiber changes randomly along a fiber link and the state-of-polarization (SOP) of an optical signal changes with environmental conditions, PMD effects on the data signal are stochastic and time varying. Therefore, any PMD compensator at a receiver must dynamically track the degrading effects due to PMD and compensate for it. Typical first-order PMD compensators consist of a polarization controller followed by a single fixed or variable differential-group-delay (DGD) element [147][182]. For the latter case,
both the polarization control and the DGD value of the compensator require accurate monitoring information.

Major previously-reported PMD monitoring techniques include either measuring the RF power or analyzing the frequency spectrum of the detected signal [147][183], or measuring the signal's degree-of-polarization (DOP) [182][184, 185]. A major problem for the above monitoring schemes is that the effective PMD value depends on both the DGD value and the principal-states-of-polarization (PSP), thereby requiring the feedback incorporating a complicated algorithm to control the polarization controller and variable DGD on a millisecond time scale [139]. Moreover, the link DGD can be ambiguous under the circumstance in which the transmitter SOP is aligned with one of the link input PSP’s, thereby causing feedback fading during the tracking procedure.

Recently, polarization scrambling of the input signal's SOP has been used to aid in PMD compensation schemes [51-53] as well as monitors that can provide either DGD or PSP information [52, 53]. A key advantage of polarization scrambling at the transmitter is that it can decouple the dependence of the monitored PMD on the input signal's SOP, thus reducing the complexity and increasing the stability of the feedback control.
In this section, we propose and demonstrate a simple technique for DGD monitoring using: (i) a 50-kHz periodic polarization scrambler at the transmitter, and (ii) the combination of a second polarization scrambler, an optical filter, and a polarization-dependent-loss (PDL) element at the receiver. The monitor signal of the instantaneous DGD is generated by the root-mean-square (RMS) value of the optical power that is fluctuating due to the polarization scrambling and subsequent transmission through the PDL element. We measure the instantaneous DGD values up to ~70 ps for both NRZ and RZ 10-Gbit/s data formats. Furthermore, as opposed with previous feedback monitoring for which the PMD or DGD monitored values incorporate both the link and compensator contributions, we demonstrate a PMD compensation configuration that uses: (i) our DGD monitor as feed-forward signal to pre-set the value of the tunable DGD element, and (ii) a traditional bit-error-rate (BER) monitor scheme that provides feedback to the polarization controller.

4.9.1 DGD monitor configuration

Figure 105 shows the DGD monitoring experimental setup. A tunable laser is externally modulated with two cascaded electro-optic (EO) modulators to achieve NRZ as well as half duty-cycle RZ intensity modulation at 10-Gb/s ($2^{23}$-1 PRBS). A 50-kHz polarization scrambler that can produce periodic polarization scrambling over a series of uncorrelated SOPs as discussed in section 4.5 is placed after the transmitter. The PMD of the link is emulated using a PMD emulator that consists
of 12 sections of polarization-maintaining (PM) fiber to emulate ~33 ps average PMD by randomly rotating the relative orientation of each PM fiber section [40].

Figure 105 Diagram of proposed DGD monitor. OF: optical filter; PD: photodiode

The concept of DGD monitoring at the receiver is described as follows. By tapping off some power of the signal at the receiver and filtering out the desired data wavelength, we monitor the instantaneous DGD using a polarization scrambler followed by a fixed 0.8-dB PDL element. The receiver polarization scrambler is used to time average the polarization coupling between the link PMD and the PDL element. If the PMD value along the link increases, then the instantaneous DOP will decrease, thus causing a decrease in the measured RMS power fluctuation value through the PDL element. This power fluctuation can be determined within one polarization-scrambling period, which could be on a sub-millisecond time scale. In this monitor configuration, locking or synchronizing of the transmitter and
receiver polarization scramblers is not necessary, and a simple polarizer can be used instead of the PDL element.

### 4.9.2 DGD monitor results

For a simple variable DGD element, Figure 106 (a) shows the measured power fluctuation variation as a function of instantaneous DGD for both NRZ and RZ 10-Gbit/s data at 1555 nm. As the DGD value changes from 0 to 70 ps, the monitor signal dynamic range due to PMD is ~20% for NRZ and ~40% for RZ. For NRZ format, consecutive "1" bits are not degraded by PMD, hence the NRZ monitored power fluctuation is less sensitive to PMD changes than is RZ.

![Figure 106 DGD monitoring results](image)

(a) Monitored DGD versus normalized power fluctuation for 10-Gb/s NRZ and RZ data formats using a variable DGD element. (b) Comparison between the DGD values of the 12-section PMD emulator given by our monitor (solid circles) and a commercial PMD analyser (dashed line).
Furthermore, in the case of all-order PMD involved, we use the 12-section PMD emulator and compare in Figure 106 (b) the instantaneous DGD of our monitor with the measured DGD using a commercial PMD analyzer. Our monitored DGD values agree well with the measured values. The difference between monitored and measured values for some points is likely due to higher-order PMD effects.

These results show the utility of this technique for a WDM system, in which an optical filter can be scanned across an entire wavelength range and the instantaneous DGD values can be rapidly acquired, thus PMD compensation can be accomplished using one DGD monitor and N traditional PMD monitors for an N-channel WDM system.

4.9.3 Feed-forward PMD compensation

In order to show the effectiveness of this DGD monitoring scheme, we further demonstrate first-order PMD compensation for a 10-Gb/s NRZ data channel. As shown in Figure 107(a), the PMD compensator is composed of a polarization controller followed by a variable DGD element. The monitored DGD is used as a feed-forward control signal to set the DGD value of the variable compensator, and the BER information at the receiver is used as the feedback signal to the polarization controller. The PMD of the link is emulated using the multi-section emulator. Compensation results down to a BER=$10^{-9}$ are shown in Figure 107(b).
The residual penalties after first-order compensation are likely due to either the higher-order PMD induced monitoring mismatch or higher-order PMD itself.

Figure 107 (a) First-order PMD compensation for a 10-Gb/s NRZ channel using monitored DGD as feed forward information. (b) Compensation results.
4.10 Simultaneous Monitoring of Both OSNR and PMD Using Polarization Techniques

A key challenge for future deployment of high-performance optical networks is that many systemic issues are not static and tend to vary with time, including: (i) temperature changes, (iii) reconfigurable optical networking, (iv) wavelength drifts, and (v) periodic repair and maintenance. These changes will cause variations in several signal-related parameters, such as optical signal-to-noise ratio (OSNR), polarization-mode-dispersion (PMD), etc. Therefore, performance monitoring of signal quality may be required in order to manage, diagnose, and repair a network. A key feature of performance monitoring would be the isolation of the specific parameter being affected, rather than a simple alarm that monitors the overall degradation of the signal.

In particular, we consider the performance monitoring of OSNR and PMD. OSNR degradation can occur due to any signal power fluctuations or changes in optical-amplifier-noise characteristics. PMD effects are stochastic, time varying, and temperature dependent. Moreover, the degradation can change rapidly due to the fiber’s nonlinear birefringence.
If there is a desire to monitor both OSNR and PMD, the typical solution would be to have two different and distinct monitoring implementations. It would be advantageous to have one module that could monitor both OSNR and PMD in a straightforward fashion. Different approaches to monitoring OSNR and PMD individually has been reported in literature [186-191]. Traditional OSNR monitoring using a high-resolution spectrum analyzer is costly, not convenient, and sometimes not accurate [187]. BER can be correlated to the OSNR only within a certain range [188]. The degree of polarization (DOP) has been used to monitor not only PMD [52, 132], but also OSNR [189, 190]. However, the OSNR and PMD (depolarisation) monitoring affect each other significantly [189, 190]. Although a PMD insensitive method to monitor OSNR is available, it requires a complicated configuration that uses precise polarization control and multiple detectors for each individual channel in the WDM systems [191].

In this section, we propose and demonstrate a simple configuration that provides both OSNR and PMD information simultaneously using polarization scrambling and orthogonal polarization detection. A narrow band optical filter is inserted in one optical path. By measuring the power fluctuation of the two orthogonal polarization channels, we can decouple optical noise and depolarisation effects, and therefore, obtain the contribution of ASE and PMD or nonlinear birefringence at the same time. Using a 10-Gb/s NRZ transmission system, we show the monitored OSNR from 18 dB to 36 dB that is not sensitive to PMD. Meanwhile, the
monitored PMD from 0-ps to 70-ps is also independent of OSNR.

4.10.1 Monitor configuration

The configuration of our proposed monitor is shown in Figure 108. Only one polarization scrambler that can scramble the input polarization state to cover the Poincare sphere is shared by many WDM channels, while previous approaches need one polarization controller (or polarization stabilizer) for each channel. After demultiplexer, the desired channel will be split into two arms using a polarization beam splitter (PBS). In one arm, there will be an optical bandpass filter with a bandwidth narrower than the channel spacing used to further reduce the effects of ASE. One photodiode is used in each arm to detect the power fluctuations. After O/E conversion, additional electrical operational amplifiers with different gain factors ($G_1$&$G_2$) are used to increase the sensitivity and provide the corresponding voltage outputs ($V_1$&$V_2$).

![Figure 108 Configuration of OSNR/PMD monitor: one PBS, two photodiodes (PD$_1$& PD$_2$) and two electrical operational amplifiers (G$_1$ & G$_2$).](image)

210
Assuming the power contributions of the signal and noise are $P_s$ and $P_N$, and that the depolarization factor along the link introduced by PMD and nonlinear birefringence is $\delta$, we then can write the power flutuations of the two arms as the following:

\[
V_{1\text{max}} = G_1[P_s(1-\delta) + 0.5P_N] \quad \text{.................................. (3.31)}
\]

\[
V_{1\text{min}} = G_1[P_s\delta + 0.5P_N] \quad \text{.................................. (3.32)}
\]

\[
V_{2\text{max}} = G_2[P_s(1-\delta) + 0.5\alpha P_N] \quad \text{............... (3.33)}
\]

\[
V_{2\text{min}} = G_2[P_s\delta + 0.5\alpha P_N] \quad \text{.................................. (3.34)}
\]

Where $\alpha$ is the noise power filtering factor of the narrow band optical filter. Using simple analysis, we can get the OSNR and the depolarization factor as

\[
S / N = \frac{G_1\overline{V}_2 - \alpha G_2\overline{V}_1}{G_2\overline{V}_1 - G_1\overline{V}_2} \quad \text{..................................(3.35)}
\]

\[
\delta = \frac{1}{2} \left[ 1 - (1-\alpha)(V_{1\text{max}} - V_{1\text{min}}) \right] \quad \text{..................................(3.36)}
\]

where $\overline{V}_i = \frac{1}{2}(V_{i\text{max}} + V_{i\text{min}})$, $i=1, 2$, is the average voltage of each arm. The monitored OSNR and PMD can be readily calculated from eqs. 3.35 and 3.36.
4.10.2 Experimental results

To show the effectiveness of our approach, we put our monitor module after a 10-Gb/s NRZ \( (2^{31}-1 \text{ PRBS}) \) transmission link. A piece of polarization-maintaining fiber (PMF) is placed along the link with a polarization controller before it to align the input polarization states with 45° to each of the two principal states of the PMF. The OSNR of the link is varied by changing the input power into an EDFA.

First we compare our approach with the traditional DOP (or equivalent DOP) method. As shown in Figure 109 (a), if only one arm is used (a polarizer followed by a photodetector), we can get the OSNR information from \( (V_{1\text{max}}-V_{1\text{min}})/V_{1\text{min}} \). Without PMD, the monitored OSNR is correlated with OSNR very well, but when PMD (here DGD \( \sim 70 \text{ ps} \)) is introduced, the monitored results will change dramatically. However, using the proposed approach (Figure 109(b)), even when large DGD is involved, we still can get the monitored OSNR results to be close to the one without PMD. The mismatch between the monitored OSNR and measured OSNR using an optical spectrum analyser (OSA) with 0.1-nm bandwidth is mainly due to the measurement accuracy of the each of the components in the equations.
Figure 109 Comparison between a previous OSNR monitor approach and the new approach (a) monitoring results using only one polarizer (i.e. one arm of the PBS), PMD has significant effect on the OSNR monitoring results; and (b) OSNR monitoring results of the proposed approach using polarization beam splitting approach (PMD-independent)

As mentioned in the previous section, we can also monitor PMD using eq 3.36. Figure 110 (a) and (b) show the PMD monitoring results. The depolarisation here is mainly due to DGD of the PMF. In Figure 110(a), we show that as the OSNR changes, the monitored depolarisation factor remains constant (<5%), and the relationship between depolarisation factor and DGD in Figure 110(b) provides an effective method to gather the PMD information of the link. For an NRZ signal, the depolarisation factor can vary from 0 to 0.2 as DGD changes from 0 ps to ~ 70 ps. As we get the OSNR and PMD simultaneously, we can separate the degrading effects due to optical noise or depolarisation, which is highly desirable for network designers and operators.
In summary, we proposed and demonstrated a novel and simple approach that can monitor two important parameters, OSNR and PMD, simultaneously using polarization scrambling and polarization beam splitting. The monitored OSNR is not sensitive to the link PMD, while the monitored PMD is also independent of OSNR. Isolation of these two parameters will enable network operators to determine separately the contribution of optical noise and depolarisation (or PMD) along the link.
4.11 Compensation of Higher-Order PMD Using Phase Modulation and Polarization Control in the Transmitter

Polarization-mode-dispersion (PMD) has emerged as a key limitation in: (i) 10-Gbit/s systems that use older legacy fiber, and (ii) 40-Gbit/s systems that use even the newest types of fiber due to the non-zero PMD of in many in-line components. Moreover, the system degrading effects caused by PMD are characterized as random stochastic processes that change with many environmental conditions [42, 128]. For such systems, dynamic PMD compensators are needed to reduce the probability of network outage.

To first order, PMD can be considered as a simple differential group delay (DGD) between the two polarization axes of the transmission fiber. Therefore, a first-order PMD compensator in the receiver employs a simple DGD element to cancel the time delay caused by the fiber. Typically, this type of PMD compensator is composed of a polarization controller, a DGD element, and a monitoring feedback loop [147]. The feedback loop is required to optimally align the incoming signal to the DGD element. After eliminating the first-order PMD effect, the system performance is essentially limited by the residual higher-order PMD [192].
Therefore, higher-order PMD compensation is desirable for more demanding high-performance systems.

Previous reports of higher-order PMD compensators have used two or more DGD sections before the receiver, with each section requiring feedback control [143, 144, 193, 141, 194]. This scenario may produce a complex software algorithm for the control feedback signal, in which the DGD elements are correlated and changing one element may perturb the optimal solution on the other element. Many regions of local power-penalty minima will appear for the feedback control loop [194, 195], thereby making optimal system performance tracking extremely difficult.

In this section, we demonstrate a new technique that compensates for higher-order PMD by using a polarization controller and phase modulator in the transmitter as well as a traditional first-order compensator in the receiver. In our method, the receiver end compensates for first-order PMD, whereas the transmitter end compensates for the higher-order PMD. These two ends are decorrelated, thereby reducing the possibility of being "trapped" in a locally-optimized state of the feedback loop. The operation is achieved by using: (i) the receiver end to compensate for first-order PMD, (ii) telemetry from the receiver to the transmitter to give information on the signal's higher-order PMD, (iii) rotating the signal SOP
at the transmitter to align with the fiber's higher-order principle axes, and (iv) producing signal chirp and higher-order compression using a phase modulator.

The effectiveness of this approach is experimentally demonstrated using a 10-Gbit/s NRZ signal and transmitting the signal over a link with an average PMD of 50 ps. Under this large PMD condition, the system power penalty can be reduced from ~ 4 dB that is achieved with a first-order compensator down to 1 dB using the proposed technique. Our simulation results show that the PMD tolerance for 40-Gb/s NRZ systems can be increased from ~ 7 ps with only first-order compensation to ~ 10 ps after higher-order compensation, according to the 30-minutes-per-year outage criterion (i.e. an outage probability of 1/18,000). Such improvement may enable doubling the transmission distance as far as PMD is concerned since PMD scales as square root of distance. The telemetry feedback response time of the high-order compensator has similar requirements as the report for first-order PSP transmission [196].

4.11.1 Concept of higher-order PMD compensation

The basic concept of higher-order PMD compensation is shown in Figure 111 using a geometric representation of PMD vectors. The first-order PMD vector along the link is denoted by $\Omega_{\text{link}}$. By applying an opposite $\Omega_{\text{comp}}$ at the receiver
by using a typical variable first-order PMD compensator [147] (i.e. a polarization controller (PC2) and a variable DGD element as shown in Figure 111(b)), we can compensate for the first-order PMD of the link. The second-order PMD vector $\Omega_\omega$ is approximately perpendicular to $\Omega_{\text{link}}$. If we align the input SOP of the signal $S_{\text{sig}}$ in parallel to the direction of $\Omega_\omega$ using a transmitter-based polarization controller (PC1 in Figure 111(b)), then: (i) the pulse can be compressed by the interaction of a given amount of phase-modulator-induced chirp with the chromatic dispersion induced by second-order PMD effects [197], and hence (ii) system degradation by higher-order PMD can be mitigated. Alternatively, the higher-order PMD compensator might use variable chromatic dispersion plus polarization control at the transmitter.

![Diagram](image_url)

Figure 111 (a) Schematic diagram illustrating the concept of proposed higher-order PMD compensation. (b) The corresponding system configuration of higher-order PMD compensation.
4.11.2 Simulation Model and Results

PMD compensation is modeled for NRZ 40-Gbit/s transmission over a fiber link that has Maxwellian PMD statistics [42]. The link contains 25 cascaded segments, with each section having a random Maxwellian distributed DGD and a randomly oriented principal-states-of polarization (PSP) that follows a uniform distribution on the Poincaré sphere [143]. Note that the link DGD distribution is exactly Maxwellian. In order to isolate the effect of PMD, we assume that chromatic dispersion and fiber nonlinearities do not accumulate along the link. The PMD-induced Q-factor penalty for all the simulated link realizations is calculated from the simulated bit-error-ratio (BER) as a function of the receiver-decision threshold, assuming that amplified spontaneous emission (ASE) is the dominant noise source.

First-order PMD compensation is achieved by compensating the link DGD to exactly zero at the signal spectral center, and the signal input state-of-polarization (SOP) is controlled by a feedback loop of two control variables to maximize the simulated eye opening. Higher-order compensation may be achieved by a pulse compression due to the interaction of chirp and the chromatic dispersion induced by second-order PMD.

The purpose of PMD compensation is to reduce the probability of inducing large system penalties or link outages that may occur infrequently in the tail of the
Maxwellian distribution. Under the assumption that an outage occurs when the Q-penalty is greater than 2.5 dB, Figure 112(a) shows the outage probability as a function of the average DGD for an ensemble of 1,000,000 independent link conditions. Similar to the technique of first-order PSP transmission [196], the feedback loop controlling the input SOP does not suffer from the existence of local optima. As a result, the outage probability calculated using a local optimization algorithm agrees with the probability calculated from a global optimization of input SOPs over the entire Poincaré sphere.

According to the criterion of the outage probability of 1/18,000, the system tolerance to PMD can be increased from 7.5 ps after first-order compensation to 10 ps after higher-order compensation. As a typical comparison, the Q-penalty distributions for both first-order and higher-order compensation at 12.5-ps average PMD are shown in Figure 112(b). Note that the above simulation results for 40 Gbit/s will generally apply for 10-Gbit/s transmission as well with a factor of 4.
Figure 112 Simulation results for 40-Gb/s NRZ transmission: (a) Probability of Q-penalties exceeding 2.5 dB. For the higher-order compensation, the open circles are calculated from the compensation based on feedback control and the solid circles are based on global optimization. (b) Outage probability as a function of Q penalty for first-order compensation and higher-order compensation with 12.5-ps average link PMD.

4.11.3 Experimental demonstration

We demonstrate our compensator using the experimental setup shown in Figure 113. The signal is generated from a wavelength-tunable external-cavity laser. The signal is NRZ modulated at 10-Gb/s ($2^{31}-1$ PRBS), with a fixed chirp introduced by a phase modulator driven by the 10-GHz clock signal. A polarization controller (PC1) is used to adjust the input SOP after the phase modulator. The PMD emulator is: (i) composed of 30-sections of polarization-maintaining (PM) fiber that are spliced at 45° angles, and (ii) provides ~50 ps of average PMD. The phase modulator generates a peak-to-peak phase change of ~ 40° that is roughly
optimized from modeling under the condition of 50-ps PMD. Higher chirp can be applied for links with lower PMD. After the PMD emulator that generates all-order PMD effects, a PMD analyzer is used to measure the instantaneous PMD condition including first-order (DGD) and second-order PMD by switching the modulated signal to the CW laser. A traditional first-order PMD compensator composed of a polarization controller (PC2) and a variable DGD element (VDGD) is used before the receiver. The DGD value of the first-order compensator is set based on the measured value of the PMD analyzer.

By turning the phase modulator on and off for the amplitude-modulated signal, we can compare the performance of first-order and second-order compensators (see Figure 114). The power penalties are compared with respect to the first-order and second-order PMD as measured from the PMD analyzer. The power penalty is also compared with the back-to-back sensitivity as measured at $10^{-9}$ BER.

Figure 113 10-Gb/s NRZ experimental setup for higher-order PMD compensation using emulator that consists of 30-section PM fiber with an average PMD of 50 ps. AM: amplitude modulator; PM: phase modulator; PC (1,2): polarization controller; VDGD: variable differential group delay (element).
Figure 114 First-order and higher-order PMD compensation results as a function of (a) DGD value (first-order PMD) (b) second-order PMD

An optical pre-amplifier before the receiver is used to increase sensitivity. We select 100 typical samples by sweeping the wavelength of the tunable laser; note that these experimental samples are not statistically distributed but are more representative of links with large second-order PMD. Throughout the experiment, we set PC1 to random polarization states for the first-order PMD compensation case. After first-order PMD compensation, the residual power penalties due to higher-order effects no longer have any correlation with link first-order PMD (Figure 114(a)), whereas the penalties are strongly correlated to the second-order link PMD (Figure 114(b)). For our experimental conditions of large second-order PMD (>1500 ps$^2$), the higher-order compensator reduced the power penalty that was achieved with a first-order compensator from ~ 4 dB down to <1 dB. Due to the pulse compression, some of the samples have negative power penalties. We
note that the phase modulator at the transmitter side can be shared by many WDM channels, or a dual-drive amplitude modulator can also be used for introducing chirp at each transmitter.

4.12 Reach Extension in 10-Gb/s Long-Haul Fiber Links with Adaptive Eye Mapping in a Si-CMOS 16-bit Transceiver IC

Intrinsic fiber impairments (dispersion, nonlinearities, and polarization effects) and accumulated optical noise introduced by optical amplifiers in the DWDM-Metro/Long-Haul, as well as Enterprise (LAN) environments, have set well-known limits on the reach/distance objectives for fiber optic transmission at bit-rates of 10 Gb/s/channel or beyond [198]. Extending the performance of practical fiber transmission systems beyond these limits has been the focus of both optical systems as well as integrated circuits research. Due to their relative simplicity, reliability and low cost, electrical optimization techniques [199] such as adaptive thresholding, feed-forward and decision feedback equalization (DFE), etc. have been proposed and demonstrated to enhance the performance of optical systems in the last few years.
For today’s reconfigurable networks, because of information path changes and environmental fluctuations, dynamic or adaptive system optimization is highly desirable. For example, optical signal-to-noise-ratio (OSNR) may vary a lot as the number of amplifiers and/or the noise figure of amplifiers change as signal propagates the networks. To date, some of the electrical methods implemented at the receiver have specially targeted different degrading effects [199-202], and some of them have no dynamic features. Specially, decision circuits employed in traditional receivers have a fixed decision threshold, and operate without a-priori information on the impairment to the signal incurred during transmission. Theoretically predicting the locus of the optimal decision point as a function of a set of system design parameters is extremely complex, if not impossible. A manual optimization (with fixed phase) is sometimes the practice if the link parameters are fixed after deployment, but practically impossible in a dynamic environment.

We note here that the sampling point quantifies both a sampling time or phase and a sampling voltage or signal strength after optical detection. It is therefore possible to expect a change in both of these quantities as a function of the link parameters, such as span length, added noise, transmitter characteristics and receiver sensitivity constraints.
Using a new CMOS decision-circuit architecture [203], we demonstrate a practical receiver with key or unique features such as: (i) low power consumption (3.3V/1.5W); (ii) adaptively estimating the optimal decision point in both phase and threshold, and (iii) a “hitless switch” allowing continuous traffic during the optimization operation. In a long-haul transmission system, we demonstrate a significant power penalty reduction (e.g. from 5 dB to ~ 2 dB at 1000 km compared with the conventional receiver that has no adaptive capabilities) and a transmission reach of up to 2000 km (~ 4.5dB power penalty) in a 10-Gb/s NRZ transmission system using adaptive sampling and equalization in the receiver decision circuit.

4.12.1 RECIEVER CONFIGURATION

The receiver is implemented in a standard 0.13 µm Si-CMOS process as a single-chip bi-directional 16:1 serializer/deserializer (commonly referred as an IC-transceiver). It consists of the following major blocks: integrated 16:1 multiplexer, 1:16 demultiplexer, adaptive dual channel clock and data recovery unit (CDR) with variable gain and bandwidth equalization and a clock multiplier unit (CMU).

At the heart of the CDR is a novel dual channel decision-circuit [200-202] with a programmable sampling point for both the phase and threshold voltage dimensions as shown in Figure 115(a). This architecture enables a statistical estimation of the BER of the received signal as a function of the sampling point. A programmable
gain amplifier and a programmable frequency response equalization unit serve as the input stages to the CDR.

Each acquisition channel and its decision circuit sample the input data at a given sampling point. The decisions of each channel are then compared to each other and the differences logged into an error counter. The error counts thus obtained can be translated into a BER or more properly, as “pseudo-BER”. In this work, we refer to this error-ratio estimate as the “relative BER” or RBER. A unique feature of our receiver is a “hitless MUX” which permits switching data traffic between either channel in order to have continuous traffic during the optimization operation.

The sampling point can be adjusted over the entire voltage and phase range of the input data eye. When both data channels are set to positions with zero RBER the optimization is considered complete. As a consequence of the optimization process, the receiver provides a two-dimensional map of the RBER vs. phase (or time) and sampled voltage, in other words, an estimate of the incoming eye-diagram. This eye-diagram can be read out of the IC if desired. Figure 115(b) shows the reported RBER eye mapping with a coarse resolution. Figure 115(c) shows the rough measured BER contour when sampling points are scanned over the data eye diagram (Note: due to the sample number limit, the contour does not exactly matching the eye).
Figure 115 Adaptive eye mapping  
(a) Operating principle (MPU: microprocessor unit), (b) typical eye diagram and (c) its corresponding BER contour plot at different sampling points across the data eye as measured by test equipment.
For comparison, we use a conventional receiver (RCVR) that consists of a limiting amplifier and a commercial CDR, which are typically used in long-reach transmission system [35]. The conventional receiver employs a single channel decision circuit with a fixed sampling point.

4.12.2 EXPERIMENTAL RESULTS

Throughout our experiment, an optical pre-amplifier is used before the receiver to increase the sensitivity. Power penalties, including OSNR degradation, are measured by comparing the receiver sensitivity at $10^{-9}$ BER to the back-to-back connection. A conventional BER test set is used to measure BER, while the measured RBER in the adaptive receiver is used for the optimization process to obtain an optimal sampling point.

OSNR is one of the most important parameters that limit system performance. In Figure 116 (a), we compare the performance of the adaptive receiver with the conventional receiver. It shows that the adaptive threshold adjustment can significantly improve the receiver performance as the OSNR changes. For example, as the OSNR varies from 33 dB to 21 dB, the power penalty of the conventional receiver increases by ~10 dB, while only ~ 5.5 dB variation is observed for the adaptive receiver. Figure 116(b) shows the BER variation across the eye diagram along the vertical direction ($S_V$) for two different OSNR conditions (28 dB and 25
dB) for the adaptive receiver. The optical power into the receiver is set to \( \sim -33 \) dBm. It clearly shows that the optimum sampling points change as the OSNR varies.

![Graph showing power penalty and OSNR comparison](image)

**Figure 116** Performance evaluation in terms of OSNR (a) comparison of power penalties as a function of OSNR. (b) BER variation across the eye diagram along the vertical direction (see Figure 115(b)) as OSNR changes for the adaptive receiver

In order to study long haul performance, we utilized a recirculating fiber loop testbed [35]. The experimental setup is shown in Figure 117(a). A single wavelength at 1555.4 nm is NRZ modulated with a \( 2^{23} - 1 \) PRBS at 10-Gb/s data rate, the modulated signal is then transmitted through the loop, which consists of three Erbium-doped fiber amplifiers (EDFAs) operating in the saturated regime, 83-km SMF and 12-km DCF. Typical BER curves for both the adaptive and conventional receivers are shown in Figure 117(b) (after \( \sim 830 \)-km transmission), with inserted eye diagram. Approximately \( \sim 1.5 \) dB and 3 dB power penalties are observed at this distance using the adaptive and conventional receiver, respectively.
The power penalty as a function of transmission distance for both the conventional receiver and the adaptive receiver are compared in Figure 117 (c). The conventional receiver shows a power penalty of ~ 5 dB at 1000-km while the adaptive receiver almost doubles the transmission distance up to ~ 2000-km with ~ 4.5 dB penalty.
In addition, we also compare the performance improvement in the presence of PMD using a first-order PMD emulator. Under the worst-case PMD scenario, the power penalties induced by 50-ps DGD are $\sim 4.8$ dB and $\sim 3.0$ dB for the conventional receiver and our adaptive receiver, respectively. This shows that the transceiver has some potential to mitigate PMD effects, although further investigation is still required.

This section (Section 4) describes related techniques for the system performance optimization, other information may be found in a series of publications by myself and my colleagues [204-210].
5 References


[114] Sumitomo Osaka Cement, Model T.MZH1.5-40.


[118] G. C. Gupta, K. Fukuchi, and T. Ogata, “Highly efficient 1Tb/s (50chx20Gb/s) 2000km RZ transmission experiment by suppressing XPM with optimized


249


