



Emerging space-time coding techniques for optical fiber transmission systems

Elie Awwad

► To cite this version:

Elie Awwad. Emerging space-time coding techniques for optical fiber transmission systems. Networking and Internet Architecture [cs.NI]. Télécom ParisTech, 2015. English. NNT : 2015ENST0004 . tel-01230644

HAL Id: tel-01230644

<https://pastel.hal.science/tel-01230644>

Submitted on 18 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Communications et Electronique »

présentée et soutenue publiquement par

Elie AWWAD

le 15 janvier 2015

**Emerging Space-Time Coding Techniques
for Optical Fiber Transmission Systems
(Techniques Émergentes de Codage Espace-Temps
pour les Systèmes de Communications Optiques)**

Directeurs de thèse : **Ghaya REKAYA-BEN OTHMAN & Yves JAOUËN**

Jury

Pr. Gabriella BOSCO, Politecnico di Torino, Italy

Pr. Juerg LEUTHOLD, ETH Zürich, Switzerland

Pr. Andrew ELLIS, Aston University, United Kingdom

Dr. Gabriel CHARLET, Alcatel-Lucent Bell Labs, France

Pr. Hikmet SARI, SUPÉLEC, France

M. Erwan PINCEMIN, Orange Labs, France

Pr. Ghaya REKAYA-BEN OTHMAN, TÉLÉCOM ParisTech, France

Pr. Yves JAOUËN, TÉLÉCOM ParisTech, France

Rapporteur

Rapporteur

Examineur

Examineur

Examineur

Invité

Directeur de thèse

Directeur de thèse

TELECOM ParisTech

Ecole de l'Institut Mines-Télécom - membre de ParisTech

46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

à Micheline, à mes parents,

Sans oublier Marie Asmar qui m'a imprégné de son savoir et de sa sagesse, je dédie ce travail à ma fiancée Micheline pour tout son amour et son courage, ainsi qu'à mes parents Johnny et Joumana pour l'éducation et l'amour qu'ils m'ont donnés.

Remerciements

Je tiens à remercier Hikmet Sari de m'avoir fait l'honneur de présider le jury de ma soutenance. Mes plus vifs remerciements s'adressent aussi à Gabriella Bosco et Juerg Leuthold pour le temps précieux qu'ils ont accordé pour lire et évaluer mon rapport de thèse. Ma gratitude va également à Gabriel Charlet, Andrew Ellis et Erwan Pincemin qui ont cordialement accepté d'être membre du jury. Erwan, merci encore une fois pour le temps que tu nous as consacré pour monter la boucle à recirculation à l'école et pour tes précieux conseils.

J'exprime ma profonde gratitude à mes directeurs de thèse Yves Jaouën et Ghaya Rekaya Ben Othman. Je n'oublierai pas leur grand soutien, ni les échanges et les discussions fructueux durant ces trois années. Leurs compétences, leur rigueur et leur dynamisme ont été déterminants dans la réalisation de mes travaux de recherche. Ghaya, Yves, je vous remercie pour votre écoute, vos conseils et surtout pour la confiance que vous m'avez accordée.

Je remercie tous les chercheurs du département Communications et Électronique de TELECOM ParisTech et d'autres universités ou laboratoires de recherche pour leurs conseils et leur soutien. Je citerai par ordre alphabétique : Jean-Claude Belfiore, Philippe Ciblat, Didier Erasme, Yann Frignac, Walid Hachem, Bruno Thedrez, Cédric Ware, Sheng Yang.

J'exprime toute ma reconnaissance aux personnes qui m'ont aidé et encouragé tout au long des années de ma thèse : Florence Besnard, Nazha Essakkaki, Bernard Cahen, Chantal Cadiat, Yvonne Bansimba, Zouina Sahnoune, Emma Schwabedissen, Deborah Dunn, Richard Bachand, Susie Heyman, Anaï Castro-Heyman, Peggy Smithhart et toute la troupe du théâtre anglais ainsi que les choristes d'ALOES.

Je n'oublierai pas l'aide et la grande amitié d'Asma Mejri, une personne formidable au grand cœur, ni les bons moments vécus avec mes collègues de bureau Julia Vinogradova et Mohamed Chaibi (l'équipe de choc !). Merci à tous mes amis thésards et post-doctorants de TELECOM ParisTech.

À Asma et à Rita Saneh, mes plus sincères remerciements pour la relecture du rapport et pour vos retours constructifs.

Mes remerciements seraient incomplets si je ne mentionnais pas mes chers amis, ma sœur Sara et toute ma famille qui ont tous contribué à rendre ces trois années encore plus belles. Qu'ils soient parvenus de tout près ou de très loin, vos anecdotes, vos mots d'encouragement et vos prières ont été vitaux !

Abstract

Research in the field of optical fiber communications is advancing at a rapid pace in order to meet the growing needs for higher data rates; a growth fueled by an increasing number of connected users and machines as well as by the emergence of more human-centered resource-hungry applications (telepresence, Internet of things, exchange of multimedia contents...). The main driving forces behind these advancements are the availability of multiple degrees of freedom in the optical fiber allowing for multiplexing more data channels, and the introduction of new digital signal processing techniques that are crucial to mitigate the impairments of the channel, hence guaranteeing a reliable communication with the highest achievable rates.

The degrees of freedom that are already used in the deployed optical transmission systems are: amplitude, phase and polarization state of the optical field, along with time and wavelength. Yet, these systems are approaching their theoretical non-linear capacity limits (around 100 Terabits/s per fiber). An extra dimension that a fiber can offer for multiplexing more information is space. Today, intensive research works are conducted to explore the benefits and costs of spatial multiplexing in optical fibers by replacing a single-core single-mode fiber with multi-mode and/or multi-core fibers to achieve the next qualitative leap in capacity. However, packing several data channels in the same medium brings with it differential impairments and crosstalk between these channels that can seriously deteriorate the performance of the transmission system. In order to guarantee a reliable communication and achieve a targeted performance over these high-capacity systems, the crosstalk should be addressed.

In this thesis, we focus on the latest optical multiple-input-multiple-output (MIMO) transmission systems based on polarization division multiplexing (PDM) and space division multiplexing (SDM). The former doubles the capacity by sending information over two orthogonal polarization states of the optical field, and the latter is a promising technique to multiply the capacity by a factor $M > 2$ by multiplexing data over many spatial modes or cores of a single fiber. In both systems, loss disparities among the channels arise due to imperfections in the used optical components (fibers, amplifiers, multiplexers...), thus resulting in non-unitary channel crosstalk. We assess the performance penalties induced by these disparities and suggest novel MIMO coding techniques: Space-Time (ST) codes, initially designed for wireless multi-antenna channels, to mitigate them.

First, for PDM coherent optical systems, we show through numerical, theoretical and experimental studies that important levels of polarization dependent loss (PDL) can be absorbed using redundancy-free Polarization-Time coded formats. We also demonstrate through our study that the best codes for the optical channel are not the same as the ones that were tailored for the wireless channel, and derive new design criteria for the PDL-impaired channel.

Second, we analyze the performance of SDM coherent transmission systems suffering from mode dependent loss (MDL). Inspired by our work on mitigating PDL, we apply redundancy-free Space-Time codes to alleviate the MDL-induced penalties in these systems. We show, for the first time, that the combination of ST coding solutions with optical solutions previously suggested to reduce, yet not completely remove MDL, can totally absorb the performance penalties in mode-multiplexed systems with $M = 3$ and 6 where high levels of modal loss disparities were observed. Later on, a complexity analysis of ST-coded solutions is carried out in order to evaluate their implementation costs and scale-up opportunities to larger optical MIMO systems $M \geq 6$. Two reduced-complexity ST schemes are hence suggested. The first solution consists in using sub-optimal decoding, thus trading a portion of the achieved coding gains with a reduced implementation complexity, while the second solution is a multi-block ST-coding approach that can be easily adapted to fit many optical MIMO schemes. Finally, apart from ST coding, we assess the value of transmit-side mode selection for MDL-impaired SDM systems as another low-complexity capacity-increasing MIMO solution.

Keywords: Multiple-Input-Multiple-Output (MIMO), Optical Fiber Communications, Polarization Division Multiplexing, Spatial Division Multiplexing, Polarization-Time Coding, Space-Time Coding, Orthogonal Frequency Division Multiplexing (OFDM).

Résumé

La recherche dans le domaine des communications sur fibres optiques avance à un rythme rapide afin de satisfaire des demandes croissantes de communications à débits élevés ; cette croissance régulière est soutenue par l'augmentation du nombre de personnes et d'objets connectés ainsi que par l'émergence d'applications centrées sur les personnes et gourmandes en bande passante (téléprésence, Internet des objets...). Les principaux moteurs de ces avancements sont la multitude de degrés de liberté offerts par la fibre permettant ainsi la transmission parallèle de plusieurs flux de données, et l'introduction de nouvelles techniques de traitement du signal cruciales pour atténuer la dégradation de ces flux engendrée par le canal de transmission optique, garantissant ainsi une communication fiable avec les débits les plus élevés possibles.

Les degrés de liberté qui sont déjà utilisés dans les systèmes de transmission optique déployés sont : l'amplitude, la phase et l'état de polarisation du champ optique, ainsi que le temps et la longueur d'onde. Pourtant, ces systèmes s'approchent de leur limite fondamentale non-linéaire de capacité (autour de 100 Térabits/s par fibre). Un degré supplémentaire que la fibre peut offrir pour multiplexer plus d'informations est la dimension spatiale. Aujourd'hui, des recherches intensives sont menées afin d'examiner les avantages et les coûts d'une solution de multiplexage spatial dans les fibres optiques qui consiste à remplacer les fibres mono-mode et mono-cœur par des fibres multi-mode et multi-cœur visant un saut qualitatif majeur en termes de capacité de transmission. Cependant, l'insertion de plusieurs flux de données dans le même canal de propagation induit également des pertes différentielles et de la diaphonie entre les flux, ce qui peut fortement réduire la qualité du système de transmission. Afin de garantir une communication fiable et cibler un certain niveau de performance minimal, l'interférence inter-canaux doit être traitée.

Dans cette thèse, nous nous concentrons sur les systèmes de transmission optique modernes à entrées multiples et sorties multiples (MIMO) basés sur un multiplexage en polarisation pour doubler la capacité de transmission en envoyant de l'information sur deux états de polarisation orthogonaux du champ optique, et en un second temps, sur les systèmes prometteurs à multiplexage spatial qui multiplieront la capacité par un facteur $M > 2$ en envoyant l'information sur plusieurs modes spatiaux de propagation et/ou dans plusieurs cœurs d'une même fibre. Dans les deux cas, des disparités de gain entre les canaux apparaissent à cause d'imperfections dans les composants optiques utilisés (fibres, amplificateurs, multiplexeurs...), aboutissant ainsi à une interférence inter-canaux non-unitaire qui détruit l'orthogonalité des flux. Nous évaluons

la dégradation de la performance provoquée par ces disparités et proposons, pour les combattre, de nouvelles techniques de codage pour les systèmes MIMO nommées “codes Spatio-Temporels” (ST), préalablement conçues pour les systèmes radiomobiles multi-antennaires.

En ce qui concerne les systèmes optiques multiplexés en polarisation, nous démontrons, à travers des simulations numériques, une étude théorique ainsi qu’une expérience de transmission, que de forts niveaux de perte dépendant de la polarisation (PDL) peuvent être absorbés en utilisant des formats de codage Polarization-Temps exempts de toute redondance. Nous prouvons aussi à travers nos études que les meilleurs codes pour le canal optique ne sont pas ceux conçus pour le canal radio-mobile, et tirons de nouvelles règles de construction de codes spécifiques aux canaux optiques affectés par la PDL.

Quant aux systèmes à multiplexage spatial, nous analysons leur performance en présence d’un effet particulièrement nuisible : la perte dépendant du mode spatial (MDL). En s’inspirant de nos travaux sur les systèmes précédents affectés par la PDL, nous appliquons des codes ST, non redondants, pour atténuer les dégradations induites par la MDL. Nous montrons, pour la première fois, que la combinaison de solutions numériques de codage ST et de solutions optiques précédemment suggérées pour réduire, sans néanmoins supprimer, la MDL, peut complètement absorber les dégradations causées dans des systèmes à $M = 3$ et 6 modes de propagation où de forts niveaux de MDL ont été observés. Par la suite, nous nous intéressons à la complexité requise par ces solutions ST afin d’évaluer leur coût d’implémentation et la possibilité de leur mise à l’échelle pour des systèmes optiques MIMO de plus grande dimension $M \geq 6$. Deux schémas ST à complexité réduite sont donc suggérés. La première solution consiste à utiliser un décodage sous-optimal sacrifiant ainsi une partie des gains de codage réalisés pour réduire la complexité d’implémentation, tandis que dans la deuxième solution, nous proposons une approche de codage multi-bloc pouvant être facilement adapté à plusieurs systèmes MIMO optiques. Enfin, outre le codage ST et dans un souci de garder la complexité du traitement numérique la plus faible possible, nous évaluons les avantages de la sélection de modes de transmission à l’émission, en tant qu’une solution d’accroissement de capacité pour les systèmes MIMO optiques affectés par la MDL.

Mots-clés : Systèmes à Entrées et Sorties Multiples (MIMO), Communications sur Fibres Optiques, Multiplexage en Polarisation, Multiplexage Spatial, Codage Espace-Temps, Codage Polarisation-Temps, Modulation OFDM.

Contents

Glossary	xiii
Notations & Symbols	xv
List of Figures	xvii
List of Tables	xxv
Introduction	1
1 Evolution of Optical Fiber Transmission Systems	7
1.1 From on-off keying to multi-level modulations	10
1.1.1 Optical transmitters	10
1.1.2 Propagation in optical fiber transmission systems	15
1.1.3 Optical receivers	22
1.2 Wavelength Division Multiplexing (WDM)	26
1.3 Polarization Division Multiplexing (PDM)	27
1.3.1 Principle of coherent PDM transmissions	28
1.3.2 Propagation equations	29
1.3.3 Polarization Mode Dispersion (PMD)	30
1.3.4 Polarization Dependent Loss (PDL)	32
1.3.5 Non-linear effects	36
1.4 Space Division Multiplexing (SDM)	36
1.4.1 Multi-core fibers	37
1.4.2 Multi-mode fibers	38
1.4.3 Linear modal crosstalk	39
1.4.4 Differential Modal Group Dispersion (DMGD)	40
1.4.5 Mode Dependent Loss (MDL)	42
1.4.6 Non-linear effects	44
1.5 Beyond 400G: WDM-SDM-PDM super-channels	44

2	Advanced Signal Processing and Coding Techniques for Optical Communications	49
2.1	Forward Error Correcting (FEC) codes	50
2.1.1	Basics of FEC	51
2.1.2	FEC in optical transmission systems	52
2.2	Basic digital equalization in 100 Gb/s single-carrier PDM-QPSK systems	54
2.2.1	Dispersion compensation & polarization de-multiplexing	54
2.2.2	Frequency offset compensation & carrier phase estimation	57
2.2.3	Transmitter-side DSP	59
2.3	OFDM: more than a modulation format	59
2.3.1	Principle	59
2.3.2	OFDM for optical communications	60
2.3.3	Capacity of linear MIMO-OFDM channels	67
2.3.4	Optical OFDM or single-carrier for MIMO schemes?	69
2.4	MIMO coding techniques	70
2.4.1	Wireless MIMO channels	71
2.4.2	Design criteria of Space-Time (ST) codes for Rayleigh fading channels .	72
2.4.3	Space-Time Block Codes	75
2.4.4	Decoders for MIMO schemes	81
3	Polarization-Time Coding for PDL Mitigation in Polarization Multiplexed Systems	87
3.1	Polarization-Time coding: preliminary results	88
3.2	Numerical study of a Polarization-Time coded optical OFDM transmission . . .	89
3.2.1	Performance of Silver, Golden and Alamouti coded schemes	91
3.2.2	Tolerance of PT-coded OFDM for non-linear effects	101
3.3	Theoretical analysis of the PT coding performance	103
3.3.1	Computing an upper bound of the pairwise error probability (PEP)	104
3.3.2	Concatenation of a FEC code and a PT code	108
3.4	Experimental validation of PT coding gains	111
3.4.1	Experimental setup	112
3.4.2	Experimental results	115
4	ST coding for MDL Mitigation in Mode Division Multiplexed (MDM) Systems	121
4.1	Description of LP modes in parabolic-index FMFs	122
4.2	MDL in MDM transmission systems	123
4.2.1	Best-case scenario: strong-coupling model	124
4.2.2	Case study 1: fiber misalignments and micro-bends	127
4.2.3	Case study 2: few mode amplifiers with mode dependent gain (MDG) . .	133

4.3	ST coding for various MDL-impaired MDM channel models	137
4.3.1	Case study 1: fiber misalignments and micro-bends	139
4.3.2	Case study 2: few mode amplifiers with mode dependent gain (MDG) . .	142
4.4	Scalability and complexity issues of ST coding	148
5	Low-Complexity Schemes for Large Optical MIMO Systems	151
5.1	Low-complexity sub-optimal decoding of ST codes	153
5.1.1	Performance of ZF decoding	153
5.1.2	Performance of ZF-DFE decoding	155
5.1.3	Performance of MMSE decoding	156
5.1.4	Performance of sub-optimal decoders in PDL-impaired PDM systems . .	157
5.2	Multi-block coding of SDM schemes	158
5.2.1	Description of multi-block schemes	158
5.2.2	Performance of ML-decoded multi-block schemes	159
5.3	Mode selection for SDM transmission systems	161
5.3.1	Motivation and principle	162
5.3.2	Application to a 10×10 MDM scheme	165
	Conclusions & Outlook	169
A	OSNR Sensitivity	173
B	Polarization of light	175
C	Matrix Operations	177
D	The Sphere Decoder	181
E	Condensed French Version	185
E.1	Introduction	185
E.1.1	Objectifs et contributions de la thèse	188
E.2	Codage Polarisation-Temps pour les systèmes de transmission optique PDM-OFDM	191
E.2.1	Modèle du système PDM-OFDM codé en polarisation et en temps . . .	193
E.2.2	Etude numérique de la compensation des effets de PDL	198
E.2.3	Validation Expérimentale de la compensation des effets de PDL	206
E.2.4	Conclusion	210
E.3	Codage Espace-Temps pour les systèmes de transmission optique SDM-OFDM	211
E.3.1	Systèmes de transmission optique multiplexés en modes spatiaux	214
E.3.2	Codage Espace-Temps et décodage en sphères	218
E.3.3	Performance des systèmes MDM codés en espace et en temps	222

CONTENTS

E.3.4	Complexité et extensibilité des systèmes MDM codés	224
E.3.5	Conclusion	229
E.4	Perspectives	230
Bibliography		233
Publications		247

Glossary

ADC	Analog-to-Digital Converter	DFB	Distributed Feedback Laser
AO(M)	Acousto Optical (Modulator)	DGD	Differential Group Delay
APol/PS	Alternating Polarization/ Polarization Switching	DMGD	Differential Modal Group Dispersion
ASE	Amplified Spontaneous Emission	DSP	Digital Signal Processing
ASIC	Application Specific Integrated Circuit	ECL	External Cavity Laser
ASK	Amplitude Shift Keying	EDFA	Erbium Doped Fiber Amplifier
AWG	Arbitrary Waveform Generator	ENoB	Effective Number of Bits
AWGN	Additive White Gaussian Noise	FDE	Frequency Domain Equalization
BCH	Boss-Chaudhuri-Hocquenghem code	FEC	Forward Error Correction/-ing
BER	Bit Error Rate	FFT	Fast Fourier Transform
BLAST	Bell Labs Space-Time Architecture	FIR	Finite Impulse Response
BPSK	Binary Phase Shift Keying	FMA	Few Mode Amplifier
CD	Chromatic Dispersion	FMF	Few Mode Fiber
CDF	Cumulative Distribution Function	FPGA	Field Programmable Gate Array
CFO	Carrier Frequency Offset	FSK	Frequency Shift Keying
CMA	Constant Modulus Algorithm	FWM	Four Wave Mixing
CMT	Coupled Mode Theory	Gb	1 Gigabit = 10^9 bits
CO-OFDM	Coherent Optical OFDM	GUE	Gaussian Unitary Ensemble
CP	Cyclic Prefix	GVD	Group Velocity Dispersion
CPE	Common Phase Error	HDD	Hard Decision Decoding
CSI	Channel State Information	ICI	Inter Carrier Interference
D(B/Q)PSK	Differential (B/Q)PSK	iFFT	inverse Fast Fourier Transform
DAC	Digital to Analog Converter	IP	Internet Protocol
DCF	Dispersion Compensated Fiber	ISI	Inter Symbol Interference
DD	Direct Detection	LDPC	Low Density Parity Check
DDO-OFDM	Direct Detection Optical OFDM	LG	Laguerre-Gauss
		LLR	Log Likelihood Ratio
		LO	Local Oscillator
		LP	Linearly Polarized
		LSE	Least Square Estimation
		MCF	Multi Core Fiber
		MDG	Mode Dependent Gain
		MDL	Mode Dependent Loss
		MDM	Mode Division Multiplexing
		MIMO	Multiple-Input Multiple-Output
		ML	Maximum Likelihood

GLOSSARY

MLSE	Maximum Likelihood Sequence Estimator	RX	Receiver
MMF	Multi Mode Fiber	RZ	Return to Zero
MMSE	Minimum Mean Square Error	SD	Sphere Decoder
MUX/DEMUX	Multiplexer/De-multiplexer	SDD	Soft Decision Decoding
MZM	Mach-Zehnder Modulator	SDM	Spatial Division Multiplexing
NLSE	Non Linear Schrödinger Equation	SFO	Sampling Frequency Offset
NLT	Non-Linear Threshold	SISO	Single-Input Single-Output
NRZ	Non Return To Zero	SM	Spatial Multiplexing
OFDM	Orthogonal Frequency Division Multiplexing	SMF	Single Mode Fiber
OOK	On-Off Keying	SNR	Signal-to-Noise Ratio
OSNR	Optical Signal-to-Noise Ratio	SOP	State Of Polarization
PAPR	Peak to Average Power Ratio	SPM	Self Phase Modulation
Pb	1 Petabit = 10^6 Gigabits	SSMF	Standard Single Mode Fiber
PDF	Probability Distribution Function	ST	Space-Time
PDG	Polarization Dependent Gain	STBC	Space-Time Block Codes
PDL	Polarization Dependent Loss	SVD	Singular Value Decomposition
PDM/Pol-Mux	Polarization Division Multiplexing	TAST	Threaded Algebraic Space-Time
PEP	Pairwise Error Probability	Tb	1 Terabit = 10^3 Gigabits
PMD	Polarization Mode Dispersion	TDE	Time Domain Equalization
PolSK	Polarization Shift Keying	TX	Transmitter
PPS	Principal Polarization State	WDM	Wavelength Division Multiplexing
PSK	Phase Shift Keying	XPM	Cross Phase Modulation
PT	Polarization-Time	XPolM	Cross Polarization Modulation
QAM	Quadrature Amplitude Modulation	ZF	Zero Forcing
QPSK	Quadrature Phase Shift Keying	ZF-DFE	ZF with Decision Feedback Equalizer
RS	Reed-Solomon		

Notations & Symbols

Notation	Definition	unit
x	scalar	
$\mathbf{x}$	vector	
$\mathbf{X}$	matrix	
$\ \cdot\ _F$	Frobenius norm	
$\lfloor \cdot \rfloor$	floor function	
$\lceil \cdot \rceil$	ceiling function	
$(\cdot)^T$	transpose	
$(\cdot)^*$	conjugate	
$(\cdot)^\dagger$	transpose conjugate	
$\mathbb{E}[\cdot]$	Expectation	
i	imaginary unit, $i^2 = -1$	
$\Im(\cdot)$	Imaginary part of a complex number	
$\Pr\{\cdot\}$	Probability	
$\Pr\{\cdot \cdot\}$	Conditional probability	
$\Re(\cdot)$	Real part of a complex number	
α_a	Fiber loss coefficient	Nepper/km
α_{dB}	Fiber loss coefficient	dB/km
A_{eff}	Fiber effective area	m ²
$\beta(\omega)$	Fiber propagation constant	rad/km
β_1	Fiber group velocity inverse	ps/km
β_2	Fiber group velocity dispersion	ps ² /km
B_{ref}	Reference spectral bandwidth	nm
c	light velocity in vacuum	m/s
C	Channel capacity	bit/s/Hz
$\Delta\tau$	Differential group delay	ps/ $\sqrt{\text{km}}$
d_{min}	minimal distance of a code or a constellation	
D	Dispersion coefficient	ps/nm/km
E_b	Energy per information bit	J
E_s	Energy per modulated symbol	J
$\hbar$	Planck constant	eV.s
γ_{NL}	Non-linearity coefficient	km ⁻¹ W ⁻¹
Γ_{dB}	Polarization dependent loss	dB
I_Q, I_I	quadrature and in-phase currents	A
k_c	Modulation spectral efficiency	
λ_c	Laser wavelength	m
L	Fiber length	m
L_{eff}	Effective length of the fiber	m
q	number of states in a constellation	
M	number of channels in an SDM system	

Notations & Symbols

$\mathbf{n}(\omega)$	Fiber refractive index	
n_{sp}	Spontaneous emission factor	W/Hz
N_{ASE}	ASE spectral density	W/Hz
N_0	Noise spectral density per real dimension	W/Hz
ω	Pulsation (angular frequency)	rad/s
Φ	Phase	rad
P_{out}	Outage probability	
τ	Photodiode responsivity	A/W
r_{FEC}	FEC rate	
r_{ST}	STBC rate	
$\mathbf{R}_\alpha$	Rotation Matrix	
R_b	Bit rate	bit/s
R_s	Symbol rate	symbol/s
t	time	s
T_s	Symbol duration	s
V_π	MZ extinction tension	V

List of Figures

1	Optical multiplexed schemes: Wavelength Division Multiplexing, Polarization Division Multiplexing, Spatial Division Multiplexing.	3
1.1	Capacity limits of standard single-mode optical fiber transmission channels. . . .	8
1.2	The evolution of transmission capacity in optical fibers as evidenced by state of the art laboratory transmission demonstrations over the years. (source: [1]). . . .	9
1.3	The general scheme of a long-distance optical transmission system.	10
1.4	Mach-Zehnder modulator: (a) schematic figure of an MZM, (b) Power transmittance with MZM biased at the quadrature point, (c) Power transmittance with MZM biased at the trough point.	12
1.5	Constellations of (a) OOK, (b) BPSK, (c) QPSK or 4-QAM having the same average energy E_S	13
1.6	Constellation diagrams of: (a) rectangular 4-QAM, (b) rectangular 8-QAM, (c) circular 8-QAM (d) rectangular 16-QAM, with the same average energy E_S	15
1.7	Single-mode fiber with a step-index profile.	16
1.8	Energy level diagram of an EDFA.	19
1.9	Inter-symbol-interference due to chromatic dispersion.	21
1.10	Characteristic bath-tub shaped BER performance as a function of the injected input power per wavelength in the optical fiber link. Extra margins $[P_{min} : P_{max}]$ may be considered rather than a single operating point P_{opt} to take into account system aging.	22
1.11	Transmitters, direct detection and coherent detection receivers for various modulation formats.	23
1.12	Schematic of the DSP blocks in a single-carrier digital coherent receiver.	25
1.13	Wavelength division multiplexing transmission scheme.	26
1.14	Architecture of dual polarization transmitter and corresponding coherent receiver.	28
1.15	General scheme of a Pol-Mux optical transmission.	29
1.16	Modeling PMD in the fiber. The incident polarization tributaries are rotated and delayed by the birefringent elements, but remain orthogonal at the output.	30

1.17	PDL element inducing unequal losses for the polarization tributaries of the input signal and breaking their orthogonality.	33
1.18	PDL effect: the reference signal without PDL propagates through a PDL element aligned with its polarization axes at left, and at an angle of $\pi/4$ with respect to its polarization axes at right.	34
1.19	The different spatial multiplexing technologies: multi-mode, few-mode fibers, and multi-core fibers where cores can be single-mode or multi-mode.	37
1.20	Spatial distributions of the first six modes given by the linearly polarized (LP) mode representation for a step-index fiber (first row) and the Laguerre-Gauss approximation for a graded-index fiber (second row). The (x,y) grid goes from $-2r_c$ to $2r_c$	38
1.21	Average and standard deviation of the OSNR penalties at $\text{BER} = 10^{-3}$ for ZF-, MMSE- and ML- decoded 3-mode transmission impaired by MDG at amplifiers (for two different gain offsets between LP_{01} and the LP_{11} modes) placed each 80 km of fibers in the weak and strong coupling regimes. Figure taken from [2]. . . .	43
1.22	Key parameters for the capacity increase of optical fiber transmission system: current single-wavelength 100G systems in black and proposed hypothetical system in green.	45
2.1	Recent progress in FEC for optical communication systems taken from [3]. AWGN stands for Additive White Gaussian Noise.	52
2.2	Time-domain equalization: (a) FIR filter structure for compensating for chromatic dispersion, (b) Butterfly structure to compensate for polarization-mode-dispersion and polarization crosstalk.	55
2.3	OFDM signal: on the left, the OFDM spectrum and a random channel frequency response; on the right, a time-domain OFDM symbol with a cyclic prefix.	60
2.4	Polarization division multiplexed OFDM transmitter with the main OFDM blocks.	62
2.5	Coherent polarization division multiplexed OFDM receiver with the main OFDM blocks.	64
2.6	OFDM frame design on the frequency-time grid showing the locations of training and synchronization symbols.	66
2.7	Multi-antenna wireless MIMO system.	72
2.8	Space-Time block coded multi-antenna wireless MIMO system.	76
2.9	Threaded Space-Time architecture for $n_t = n_r = \{3,6\}$	79
2.10	Constellations of STBCs using 4-QAM symbols: Alamouti, Golden and Silver codes.	80

2.11	Golden and Silver code performance on 2×2 Rayleigh fading channel with ML decoding. The performance of some sub-optimal decoders of the uncoded and the Golden-coded schemes is also shown. The AWGN-channel performance is given as a reference.	82
3.1	General scheme of a PDM-OFDM transmission system with PT coding.	90
3.2	The three investigated long-haul optical links	92
3.3	Average BER curves of 4-QAM (No coding), Silver, Golden and Alamouti coding schemes for different constant Γ_{dB} in a 20×100 km link.	93
3.4	BER dependence on the parameters θ and ν of the unitary matrix $\mathbf{V}_k$ in the SVD decomposition of $\mathbf{H}_k$ for two levels of PDL at SNR = 10 dB.	95
3.5	Simulated and theoretical probability distribution functions of the overall PDL Γ_{dB} of a concatenation of low-PDL and birefringent elements.	96
3.6	Frequency-dependence of PDL due to PMD in the optical fiber link.	98
3.7	In '×' marks, the simulated distribution of Γ_{dB} and the corresponding theoretical Maxwellian distribution of mean 6 dB in red. In '+' marks, the simulated distribution of the effective PDL of $\mathbf{H}_k$ and the corresponding theoretical Maxwellian of mean 4 dB in blue.	100
3.8	Simulation model of a long-haul PDM optical fiber link using the Split-Step Fourier method to include non-linear effects in the SMF fiber.	101
3.9	Performance of PT-coded PDM-OFDM system in presence of Kerr non-linear effects.	102
3.10	Plot of the 0^{th} order modified Bessel function of the first kind.	105
3.11	Bit Error Rate as a function of the SNR_{bit} for the uncoded scheme and the Silver code, with or without FEC: (a) Hard decision decoding (HDD), (b) Soft decision decoding (SDD). The simulated FEC is a BCH(63,45) code.	109
3.12	Experimental setup. (ECL: External Cavity Laser, AWG: Arbitrary Waveform Generator, MZM: Mach-Zehnder Modulator, MUX: Multiplexer, AO: Acousto-Optical Modulator, PS: Polarization Scrambler, OBPF: Optical Band-Pass Filter, ASE: Accumulated Spontaneous Emission source, LO: Local Oscillator, OSA: Optical Spectrum Analyzer, OSC: Tektronix 50GS/s Oscilloscope).	112
3.13	Acousto-optic switches for controlling the loop.	114
3.14	BER curves versus $OSNR_{0.1nm}$ with PDL = 0, 3, 6 dB for the different schemes: no coding, Silver, Golden and Alamouti codes.	115
3.15	Q -penalty versus PDL at $OSNR_{0.1nm} = 13$ dB with regard to a PDL-free scenario.	115
3.16	Comparison of BER evolution versus launched input power after 1000 km for the Silver-coded and QPSK schemes (PDL at the transmitter = 0, 3, 6 dB).	116

3.17	Experimental results.	118
4.1	Various graded-index profiles with a refractive index $n(R) = n_{co} (1 - 2\Delta R^2)^{1/\alpha}$ where $\Delta = (n_{co} - n_{cl})/n_{co} \ll 1$ and $R = r/r_c$ the normalized radius. Special cases: $\alpha = 2$ parabolic index, $\alpha = \infty$ step-index.	123
4.2	Spatial distributions of the LP_{01} , LP_{11a} and LP_{11b} , LP_{21a} and LP_{21b} , and LP_{02} modes given by the Laguerre-Gauss approximation: 3D view on the left, and 2D view of (x, y) plane on the right.	124
4.3	Probability distribution functions of the modal gains for different number of sections $K = \{32, 64, 128, 256\}$ with $\xi = 1$ in units of logarithm of power gain compared to the eigenvalues of the equivalent zero-trace GUE.	126
4.4	Probability distribution functions of the maximum MDL for $M = \{2, 4, 6, 8, 64\}$	126
4.5	SDM channel model impaired by fiber misalignments and micro-bends.	127
4.6	SDM channel model impaired by fiber misalignments and micro-bends with periodic mode scrambling (MS).	128
4.7	Mode-multiplexed OFDM optical transmission system.	129
4.8	PDF of MDL defined as the ratio between the highest and the lowest singular values of the 3×3 channel for different misalignment levels and various mode scrambling scenarios ($K = 300$).	130
4.9	BER curves of the ML-decoded 3×3 MDM system (spectral efficiency = 6 bits/cu).	131
4.10	PDFs of modal gains of the 3×3 channel for a misalignment of 4% and various mode scrambling scenarios.	132
4.11	Mode-multiplexed OFDM optical transmission system.	135
4.12	Simulated probability distribution functions of MDL in the investigated 6×6 MDM system in different coupling and scrambling scenarios (8 spans, 2 dB of MDL per span).	136
4.13	CDF of the capacity of the 6×6 SDM system in the different coupling and scrambling scenarios at SNR = 10 dB.	136
4.14	Performance in terms of average BER versus SNR of 3×3 (a,c,e) and 6×6 (b,d,f) MDM systems obtained through Monte Carlo simulations (8 spans of graded-index parabolic profile fibers with a core radius $r_c = 6 \mu\text{m}$ (resp. $r_c = 8.7 \mu\text{m}$) and a numerical aperture $NA = 0.205$ at a wavelength $\lambda = 1550 \text{ nm}$, FMAs with a maximum modal gain offset of 2 dB.	138
4.15	Mode-multiplexed Space-Time coded OFDM optical transmission system.	139
4.16	BER curves of the 3×3 TAST-coded SDM system ($K = 300$) with a spectral efficiency of 6 bits/c.u. and using ML decoding.	141

4.17	BER curves and MDL statistics of the 3×3 TAST-coded SDM system ($K = 300$) with $\sigma_{x,y} = 5\%r_c$ and using ML decoding.	142
4.18	Performance in terms of average BER versus SNR of 3×3 (a,c,e) and 6×6 (b,d,f) ST-coded MDM systems obtained through Monte Carlo simulations (8 spans of graded-index parabolic profile fibers with a core radius $r_c = 6 \mu\text{m}$ (resp. $r_c = 8.7 \mu\text{m}$) and a numerical aperture $NA = 0.205$ at a wavelength $\lambda = 1550 \text{ nm}$, FMAs with a maximum modal gain offset of 2 dB.	144
4.19	Evolution of the average accumulated MDL with the number of spans L for various coupling strengths in absence of mode scramblers (2 dB MDL per span).	146
4.20	Ratio of mode scramblers required to optimally reduce MDL for various coupling strengths determined by σ	147
4.21	Ratio of mode scramblers required to optimally reduce MDL for various coupling strengths determined by σ in the asymptotic regime.	147
5.1	Performance in terms of average BER versus SNR of ZF-decoded 6×6 ST schemes with mode scrambling, obtained through Monte Carlo simulations.	154
5.2	Performance in terms of average BER versus SNR of ZF-DFE-decoded 6×6 ST schemes with mode scrambling, obtained through Monte Carlo simulations. . . .	156
5.3	Performance in terms of average BER versus SNR of MMSE-decoded 6×6 ST schemes with mode scrambling, obtained through Monte Carlo simulations. . . .	156
5.4	Performance in terms of average BER versus SNR of sub-optimal decoding of PDL-impaired 2×2 ST schemes for $\Gamma_{\text{dB}} = \{3, 6\}$ dB.	157
5.5	Representation of codewords over modes (space) and time for the investigated schemes: (a) Single block with a 6×6 code, (b) Two blocks with 3×3 codes, (c) Three blocks with 2×2 codes, (d) Uncoded spatial multiplexing.	158
5.6	Different ST-coded SDM transmission schemes with the simulated optical channel (MS: mode scrambler).	159
5.7	Performance in terms of average BER versus SNR of multi-block ST schemes for 6×6 MDM systems with mode scrambling, obtained through Monte Carlo simulations.	160
5.8	PDFs of gains/losses of the 6 modes after propagating in a SDM system made of $K = 300$ misaligned sections with two different misalignment levels.	163
5.9	Average magnitudes of channel coefficients of a 6-mode SDM system impaired by MDG at few-mode amplifiers in two coupling scenarios (The grid represents the channel matrix $\mathbf{H}$ and $1 = LP_{01}$, $2 = LP_{11a}$, $3 = LP_{11b}$, $4 = LP_{21a}$, $5 = LP_{21b}$, $6 = LP_{02}$).	163
5.10	SDM transmission system with mode selection.	164

5.11	Performance in terms of average BER versus SNR of 3-mode fiber and 6-mode fiber using M_T transmit modes and M_R modes at the receiver side, with a spectral efficiency of 6 bits/cu.	165
5.12	Performance in terms of average BER versus SNR of 6-mode fiber and 10-mode fiber using M_T transmit modes and M_R modes at the receiver side, with a spectral efficiency of 12 bits/cu.	166
A.1	SNR sensitivities of various modulation formats: average BER versus SNR = $E_S/2N_0$ over an AWGN channel.	173
B.1	The polarization ellipse completely describes the polarization state: A is the major semi-axis, B the minor semi-axis, ψ the orientation of the ellipse (tilt or azimuthal angle), $\chi = \arctan(\frac{B}{A})$ the ellipticity angle, and $h = \pm 1$ the sense of rotation (Image source: Wikimedia Commons).	175
B.2	Representation of the Poincaré sphere with the three Stokes parameter $S_{1 \rightarrow 3}$ (Image source: Wikimedia Commons).	176
D.1	An illustration of a sphere centered at the received symbol in a 2-dimensional lattice.	181
E.1	Le multiplexage dans les systèmes de transmission optique : Wavelength Division Multiplexing (WDM), Polarization Division Multiplexing (PDM), Spatial Division Multiplexing (SDM).	187
E.2	Schéma général d'une transmission PolMux OFDM avec codage Polarisation-Temps.	193
E.3	Structure considérée du lien optique de longue distance.	195
E.4	Performance des codes Polarisation-Temps obtenus par simulations Monte-Carlo.	200
E.5	TEB en fonction du SNR_{bit} pour le cas sans codage MIMO et le cas du Silver code, sans et avec codage FEC : (a) Décodage à décision dure HDD), (b) Décodage à décision souple (SDD). Le code FEC simulé est un code BCH(63,45).	201
E.6	TEB en fonction du SNR_{bit} pour le Silver code, obtenu par simulations Monte-Carlo.	205
E.7	Schéma expérimental. (ECL : External Cavity Laser, AWG : Arbitrary Waveform Generator, MUX : Multiplexeur, AO : Acousto-Optical Modulator, PS : Polarization Scrambler, OBPF : Optical Band-Pass Filter, ASE : Accumulated Spontaneous Emission source, LO : Local Oscillator, OSA : Optical Spectrum Analyzer, OSC : Tektronix 50GS/s Oscilloscope).	207
E.8	Évolution du TEB en fonction de la puissance injectée après $5 \times 200\text{km}$ pour le cas non-codé et pour le Silver code, à trois niveaux de PDL différents : 0, 3 et 6dB.	209

E.9	Distributions des facteurs Q après une transmission de $5 \times 200\text{km}$ à $P_{in} = -3\text{dBm}$ ($\text{OSNR}_{0.1nm} = 12\text{dB}$). Insertion : Distributions théorique et expérimentale de la PDL.	210
E.10	Système de transmission optique OFDM multiplexé en modes, et codé en espace et en temps.	216
E.11	Distributions de probabilité de la MDL simulées pour différents scénarios de couplage et de brouillage de modes dans un système MDM 6×6 (8 tronçons, 2dB de MDL par tronçons).	218
E.12	TEB moyens en fonction du SNR pour un système MDM 3×3 (a,c,e) et 6×6 (b,d,f) obtenus par des simulations Monte-Carlo (8 tronçons de fibres à gradient d'indice avec un profil d'indice parabolique et un rayon de cœur de $r_c = 6\mu\text{m}$ (resp. $r_c = 8.7\mu\text{m}$) et une ouverture numérique $NA = 0.205$ à $\lambda = 1550\text{nm}$, FMAs avec décalage maximal de gain modal de 2dB).	223
E.13	TEB moyens en fonction du SNR du décodage ZF-DFE des schémas codés en TAST 6×6 avec brouillage de modes.	227
E.14	Quatre configuration de mots de codes répartis en espace et en temps : (a) code 6×6 en un bloc unique, (b) deux blocs de codes 3×3 , (c) trois blocs de codes 2×2 , (d) Simple multiplexage spatial.	228
E.15	TEB moyens en fonction du SNR des schémas de codage multi-bloc du système MDM à 6 modes avec brouillage de modes.	229

List of Tables

3.1	Simulation parameters related to the optical channel on the left and to the OFDM modulation format on the right	90
3.2	d_{min}^2 for different coding schemes at different PDL values	107
4.1	Loss and crosstalk levels in dB between the LP_{01} , LP_{11a} and LP_{11b} modes at the output of three different misalignment scenarios.	129
4.2	Comparison of the complexity of different ML decoders of $N \times N$ MIMO systems.	148
5.1	Comparison of the complexity of different decoders of $N \times N$ MIMO systems.	153
5.2	Parameters of the graded-index fibers used in the simulation.	165
E.1	d_{min}^2 des différents schémas de codage aux différentes valeurs de PDL	204
E.2	Comparaison de la complexité de décodage par symbole de différents décodeurs d'un système MIMO de taille $N \times N$	226

Introduction

From access and metropolitan networks to long-haul terrestrial and transoceanic links, optical fibers are the unrivaled carriers of data traffic over the world, mainly owing to their capability to transport the modulated light over long distances with low attenuation and over a large bandwidth. During the last twenty years, optical-fiber based communication systems have known important and steady evolution in order to meet the increasing demand for higher capacities in environments that are becoming, more and more, information-driven. The ever-growing number of users and machines connected to the Internet urges the need to increase the data-carrying capacity of every single optical fiber in the backbone of the global telecommunication system.

According to the 2014 Cisco Visual Networking Index (VNI) report [4], global IP traffic has increased fivefold over the past 5 years, and will increase threefold over the next 5 years. In particular, busy-hour Internet traffic (the busiest 60-minute period in a day) is growing more rapidly than average Internet traffic and will reach 1 Petabits per second (Pb/s) in 2018 while the average Internet traffic will reach 311 Terabits per second. Moreover, metropolitan traffic in the area of a metropolis or a country (scaling from 100 to 1000 km) is expected to surpass long-haul traffic (covering longer distances ≥ 3000 km) in 2015, accounting for 62% of total IP traffic by 2018. This is partially due to the widespread installation of regional content delivery networks offering video on demand services. The traffic growth rates will also be boosted by the IP traffic originating from newly connected devices such as smart TVs, tablets, smartphones and other mobile devices and machine-2-machine modules.

In order to respond to these growing and urgent demands for bit rates, research in optical-fiber communications introduced a number of innovative solutions that led to the development of high bitrate optical transmission systems. This evolution involved the exploration and use of numerous degrees of freedom in a standard single mode fiber (SSMF):

1. **Amplitude & Time :** First commercial optical transmission systems consisted in intensity modulation (on-off keying or OOK) of light along with direct detection at the receiver. Without any advanced processing at the receiver, the achieved bit rates that guaranteed an acceptable signal quality reached 10 Gigabits per second (Gb/s) over important distances. Increasing the bit rates to 40 and 100 Gb/s was not possible over the same distances due

to the higher optical signal-to-noise ratio (OSNR) requirements and a decreased robustness with respect to chromatic dispersion and polarization mode dispersion (PMD).

2. **Wavelength :** The impressive development of optical amplifiers enabled the transmission of independently modulated wavelengths over thousands of kilometers without opto-electrical regeneration, known as wavelength division multiplexing (WDM) illustrated in Fig. 1a. Modern WDM systems can transmit up to 80 channels in parallel over a 4 THz spectral window partitioned into 50 GHz slots. Each wavelength can be modulated at 10, 40 and currently 100 Gb/s yielding total capacities of more than 1 Terabit per second (Tb/s) over a single fiber.
3. **Phase :** Although differential binary phase shift keying (DBPSK) with direct detection were considered to enhance the OSNR sensitivity compared to OOK, a disruptive solution offering an increased capacity and robustness with respect to the linear impairments of the optical channel was needed. The well-known coherent detection reemerged as a key tool for high-capacity optical links, empowered by the innovations in high-speed electronic devices and digital signal processing. By giving access to both amplitude and phase of the optical field, coherent receivers allowed to increase both the bitrate and the transmission distance of the optical transmissions thanks to the equalization of the impairments of the optical channel, mainly dispersion effects, in the electrical domain using advanced DSP techniques. Net bit rates of 40 Gb/s per wavelength per fiber were achieved over thousands of kilometers using quadrature phase shift keying modulation (QPSK).
4. **Polarization :** Coherent detection can also give access to the polarization state of the electrical field of the optical wave. Any electrical field can be expressed in a vector form and arbitrarily divided into two orthogonal polarization states (also called polarization modes). The next improvement in the data-carrying capacity that allowed to achieve 100 Gb/s per wavelength per fiber, was enabled by polarization division multiplexing (PDM or Pol-Mux) illustrated in Fig. 1b. This technique consists in sending independent data symbols on two orthogonal polarization states of the optical signal, hence doubling the spectral efficiency of the optical transmission while being implemented on the legacy infrastructure of the previous single-polarization systems. Indeed, over the same spectral bandwidth, with the same OSNR requirements and the same sampling rates at transceivers (e.g. integrated circuits processing 28 GSymbols/s or GBaud), a total bitrate of $2 \text{ pol} \times 2 \text{ bits/pol} \times 28 = 112 \text{ Gb/s}$ is reached by a PDM-QPSK system.

100 Gb/s optical transceivers are already available and research is currently exploring low cost and energy efficient solutions to provide, in the near future, 400 Gb/s per wavelength per fiber on the

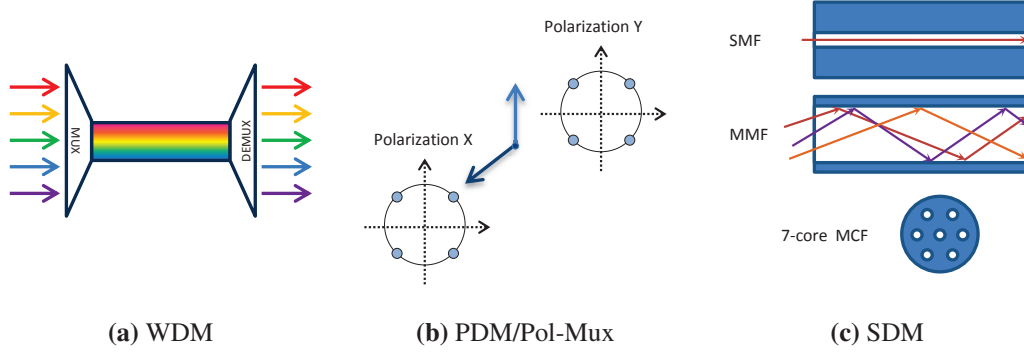


Figure 1: Optical multiplexed schemes: Wavelength Division Multiplexing, Polarization Division Multiplexing, Spatial Division Multiplexing.

currently deployed infrastructure. 16-QAM modulation, orthogonal frequency division multiplexing (OFDM) or other multi-carrier formats, and advanced forward error correction (FEC) codes are promising candidates for achieving this target. However, the increased sensitivity of higher-order modulation formats to noise and channel impairments will require higher OSNR levels and the implementation of more advanced FEC decoding algorithms will require larger integrated circuits dissipating more power. On top of that, the capacity of SSMF-based optical links has a fundamental limit due to the non-linear effects in the fiber (mainly the Kerr effect: a change in the refractive index of the fiber with the intensity of the propagating field). Hence, the capacity will not be able to exceed a certain peak when increasing the injected power in the fiber, a peak that decreases with the system length [1, 5, 6]. Indeed, while the maximum achievable capacities in the latest transmission experiments double every year and a half, the maximum capacity distance product doubles every five years [7]. Thus, a disruptive solution is needed to keep up with the growing demands for capacity.

Space is a degree of freedom which is not yet used in the fiber for multiplexing, and is being intensively investigated today for the next leap in capacity. Space division multiplexing (SDM) can be implemented with multi-mode fibers (MMFs) having a large core diameter allowing the propagation of more than one mode in contrast to the currently deployed single mode fibers (SMFs) as illustrated in Fig. 1c. SDM can also be realized through the use of multi-core fibers (MCFs) that have more than one core in the same cladding. The available spatial modes or cores form an orthogonal set of channels over which independent data symbols can be multiplexed.

A WDM/SDM/PDM optical transmission system is strongly imposing itself as a future Petabit-achieving solution for short-reach, metropolitan, long-haul or any point-to-point ultra-high-capacity link. Besides the capacity increase per fiber, an important advantage of deploying SDM-based, instead of more SMF-based solutions, is the possibility of offering integrated components (a single amplifier for all channels, integrated transceivers...). Integration will avoid multiplying the cost and the energy consumption of the transmission system by the same capacity increase factor.

However, a tight packing of channels in the same medium brings with it differential impairments and crosstalk between channels that can seriously deteriorate the performance of the system. In order to guarantee a reliable communication and achieve a targeted performance over these high-capacity multiplexed systems, the crosstalk should be addressed.

Goals and contributions of the thesis

In this thesis, we propose new digital signal processing techniques to improve the performance of high-capacity multiplexing solutions for today's and tomorrow's optical transmission systems impaired by non-unitary crosstalk effects. Non-unitary means that it causes a loss of orthogonality and/or loss of energy between the channels over which data is multiplexed. While currently used coherent transceivers are equipped with linear equalizers that address dispersive effects (chromatic dispersion and PMD) that are by nature unitary, they do not deal with non-unitary effects. In particular, we focus on mitigating the linear capacity-limiting polarization (resp. mode) dependent loss in PDM (resp. SDM) systems with coding techniques designed for multiple-input multiple-output (MIMO) channels.

In 2010, Mumtaz *et al.* [8] showed that a family of codes called Space-Time (ST) codes, initially designed for wireless multi-antenna systems, efficiently mitigate polarization dependent loss (PDL) effects in PDM systems. Numerical simulations and a preliminary experimental demonstration were conducted. The obtained performance on the optical channel showed significant differences in terms of achieved bit error rates with respect to the wireless Rayleigh fading channel. These results initiated this thesis work that started in October 2011 and aimed to further develop the obtained results and suggest ST coding solutions for SDM transmission systems.

In this thesis project, a simulation framework was first built where a PDM-OFDM transmission system is defined and the resistance of ST codes to linear and non-linear effects of the optical channel was investigated. OFDM was chosen as the preferred format for a low-complexity implementation of ST codes. Then, using a global optical channel model impaired by PDL, a theoretical analysis was conducted by computing an upper bound of the error probability of a transmission over this channel, in order to explain the performance of different codes and to design appropriate ST codes for PDL mitigation. Furthermore, knowing that most current optical transmission systems use FEC codes to improve the performance or the reach, we added a FEC module to our scheme and observed a summation of the FEC and the ST coding gains. This unexpected summation was also explained using the error probability analysis of the PDM channel.

Following these studies, we completed our work with an experimental validation of PDL mitigation. To this end, an experimental bench consisting of a recirculating fiber loop with in-line PDL, was arranged in the laboratories of TÉLÉCOM ParisTech. We demonstrated, for the first time, the mitigation of PDL in a 1000km-long PDM optical link using ST codes.

Later on, having thoroughly analyzed PDL mitigation, we were interested in combating mode dependent loss (MDL) in mode division multiplexed (MDM) systems, a sub-category of SDM. MDL can seriously deteriorate the performance of MDM transmission systems. Different combinations of loss disparities can arise in these systems where a maximum number of data-carrying modes is desired in order to expand the capacity. Recent advancements in optical fiber technology led to the fabrication of fibers supporting a small number of modes, usually 3 or 6 modes, and having interesting properties [9, 10]. These fibers are called few-mode fibers (FMFs). Few-mode optical components such as optical amplifiers were also designed and optimized. Scaling these technologies to a larger number of modes is still a real challenge. In our work, we considered two main sources of MDL resulting from imperfections in FMFs (bends, splices, connections...) and gain offsets in few-mode optical amplifiers, periodically inserted in a link to compensate the fiber attenuation. In both cases, we succeeded to completely mitigate important levels of MDL using ST codes as standalone solutions or complementary solutions to some optical-based MDL reducing techniques.

Finally, we focused on the complexity and the scalability of ST solutions. A major downside of ST codes is the rapid increase of their decoding complexity with the dimensions of the MIMO system. Trading a part of the achieved coding gains with an important reduction of the decoding complexity, we succeeded to define two near-optimal ST coding schemes that can be easily scaled up for large MIMO systems. Moreover, apart from ST coding, and taking into account the particularities of some modal crosstalk effects in multi-mode fibers and associated optical components, we have proposed mode selection techniques for large MDM systems where the best modes can be selected for data multiplexing, reducing thus the DSP complexity at the transceivers while guaranteeing an improved transmission quality.

Thesis Outline

Accordingly, the thesis is organized as follows: in chapter 1, we briefly review the evolution of optical transmission systems from basic time division multiplexing and amplitude modulation to modern optical MIMO transmission systems. Then, we focus on the linear and non-linear impairments affecting PDM and SDM optical links, starting with the common impairments and ending with the particular ones in each scheme. Among the latter, we pinpoint the sources and effects of capacity-limiting linear impairments known as polarization dependent loss in PDM systems and mode dependent loss in MDM systems.

Chapter 2 is dedicated to the description of the digital signal processing tools that will be applied to the optical MIMO systems. First, the principles of FEC coding and its application in optical transmission systems are recalled, followed by the principles of time-domain and frequency-domain equalization of dispersive effects with a special focus on OFDM. Then, MIMO coding

techniques are introduced with a review of different ST coding families and their corresponding decoding strategies.

Chapter 3 introduces the concept of polarization-time (PT) coding for PDL mitigation. We recall previous works and evaluate the performance of a PT-coded OFDM PDM system in the linear and non-linear regimes. Next, we derive an upper bound of the error probability of an optical channel with PDL in order to explain the performance of the different PT codes. We also define the design criterion of codes that optimally mitigate PDL. Later on, an experimental setup is described and used to validate the observed coding gains and demonstrate the mitigation of in-line PDL.

In chapter 4, the latest works on MDL mitigation are recalled, showing that more sophisticated signal processing is needed to absorb the induced penalties. Then, ST codes are applied to 3-mode and 6-mode MDM systems where low dispersion fibers and optical amplifiers are installed. Two simulation scenarios are investigated where MDL arises from imperfections in the fiber itself or from gain offsets at optical amplifiers. Finally, we analyze the DSP complexity of the suggested ST coding solutions.

Chapter 5 deals with the reduction of the DSP complexity of the proposed optical MIMO solutions. First, sub-optimal low-complexity decoders are suggested and their performance compared to the one obtained with an optimal decoding. Then, a concept of multi-block ST coding is presented and applied for MDL mitigation. Finally, mode selection is proposed as another complexity reducing technique for large SDM systems.

Chapter 1

Evolution of Optical Fiber Transmission Systems

Optical fiber communication systems consist in modulating light, emitted by optical sources such as lasers, with an electrical data-carrying signal. The modulated optical signal then propagates in the fiber which is a transparent cylindrical waveguide principally made of silica. Light is trapped in the core of the waveguide by internal reflections between the core and the cladding that has a slightly different refraction index. At the end of the link, the signal is detected using photo-receivers that convert the carried data back into the electrical domain. Propagation distances can vary from few meters in local networks (homes, enterprises,...) to thousands of kilometers in transoceanic links empowered by groundbreaking optical amplification techniques. The key advantages of optical fibers, introduced in the 1970s as a transmission medium, are their low attenuation and wide spectral bandwidth.

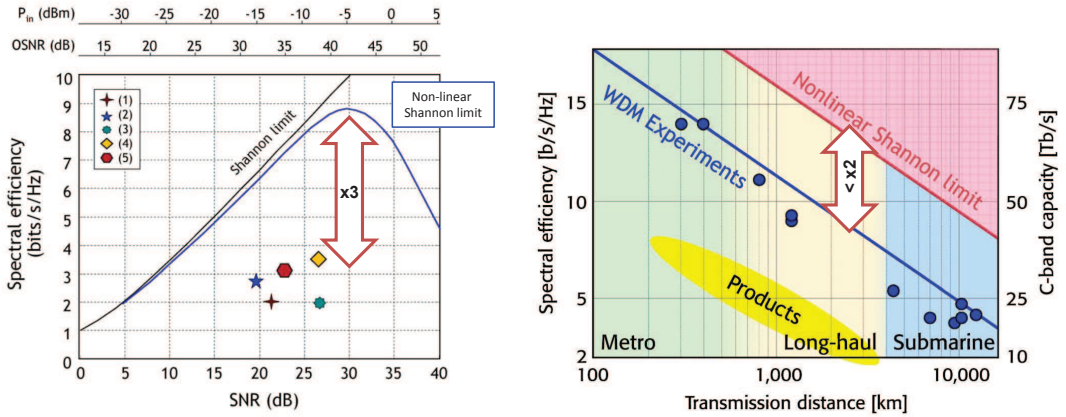
In the last twenty years, wavelength division multiplexing (WDM), coherent detection and polarization division multiplexing (PDM) along with advanced signal processing techniques marked important milestones in the evolution of optical fiber transmission systems, enhancing both, their capacity and their maximum reach. While simple modulation formats such as OOK or BPSK along with direct detection provided bit rates ranging from 2.5 to 40 Gb/s, the highest being achieved using strong FEC codes, the reach of these systems is limited by dispersive effects (CD and PMD) that induce pulse broadening.

Later on, coherent detection provided access to all the information on the complex envelope of the optical field (amplitude, phase, polarization state) allowing the compensation of dispersive effects at the receiver and the doubling of the capacity through polarization multiplexing, hence enabling 40 to 100 Gb/s rates per wavelength over the same legacy optical infrastructure (fibers and amplifiers). For instance, a single-wavelength PDM system with a net bitrate of 100 Gb/s can be implemented with a symbol rate of 28 GSymbols/s (or GBaud) and an optical bandwidth

1. EVOLUTION OF OPTICAL FIBER TRANSMISSION SYSTEMS

of 50 GHz using QPSK data symbols. An upgrade to 400 Gb/s per wavelength is expected in the near future and can be obtained with the same symbol rate using a “super-channel” made of two adjacent 50 GHz optical channels and using 16-QAM symbols, or a single 100 GHz optical channel and a symbol rate of 56 GBaud with 16-QAM symbols. For all these systems, WDM can scale up the capacity by the number of wavelengths transmitted in the fiber.

Today, around 10 Tb/s WDM systems are commercially available and 100 Tb/s are demonstrated in research laboratories. However, coherent WDM-PDM systems are approaching their capacity limits [6]. The most important difference between optical fiber channels and other transmission media is the presence of non-linear effects in the fiber. Light is confined within a very small volume: a fiber core with a radius of few micrometers, hence increasing the injected power of the optical signal would lead to non-linear distortions that will accumulate over hundreds of kilometers. This non-linear property has important consequences on the capacity of the channel as demonstrated in [5, 6]. While the capacity of a linear channel impaired by additive Gaussian noise increases monotonically with the transmitted power [11, Chap.6], the impact of non-linear distortions in the fiber can grow at a faster rate than the gain brought by the high signal power, turning the capacity into a non-monotonic function of the transmitted power as can be seen in Fig. 1.1a.



(a) Spectral efficiency results for record experiments from [6] compared to the non-linear capacity limit estimate curve for a 500-km transmission.

(b) Achieved spectral efficiency and transmission distances in WDM experiments compared to the nonlinear Shannon limit, taken from [12].

Figure 1.1: Capacity limits of standard single-mode optical fiber transmission channels.

Record experimental capacities are about a factor of two to three from the theoretical capacity limit estimate in the few-thousand kilometer transmission regime. In Fig. 1.1b, we notice that the capacity limit also known as the non-linear Shannon limit decreases with the transmission distance. This is mainly due to the fact that the periodically amplified optical signal (usually every 60 to 120 km) is accumulating more non-linear distortions. Even though the optical fiber can be

enhanced to increase its capacity, for instance by further reducing its loss or increasing its core area, this capacity gain is not enough to cope with the rising demands for higher bit rates [1]. Therefore, a new degree of freedom is currently investigated in order to provide the next leap in capacity.

This new degree of freedom is space. All currently deployed long-distance optical transmission systems use silica fibers with a small core radius where only one propagation mode, defining how a wave is distributed while propagating in the fiber, can exist. These fibers are called single-mode fibers (SMFs). Cheaper fibers with a larger core, known as multi-mode fibers (MMFs) where up to 100 modes can propagate at the same time are used for distances of few meters in local networks. However, these MMFs are not optimized for data multiplexing. The modes propagate at disparate velocities and experience various losses. Recently, advancements in fiber design led to a new category of fibers called “few-mode fibers” (FMFs) that support a small number of modes with interesting properties [10], making them suitable for long-distance communications. Moreover, fibers with multiple cores in the same cladding were also suggested [13]. These promising innovations can allow for overcoming the NL Shannon-limit of SMFs through spatial division multiplexing (SDM). In Fig. 1.2 showing the evolution of the capacity of optical transmission systems over the years, the red data points correspond to the capacities achieved by SDM solutions recently demonstrated and presented in post-deadline papers at the past European Conference on Optical communications (ECOC) editions starting from 2010. SDM systems are able to offer Petabit/s (Pb/s) capacities and are hence candidates for the next key technological breakthrough in optical fiber transmission systems, just like low-loss fibers, optical amplification, WDM and more recently high-spectral efficiency modulation via DSP-enabled PDM coherent systems.

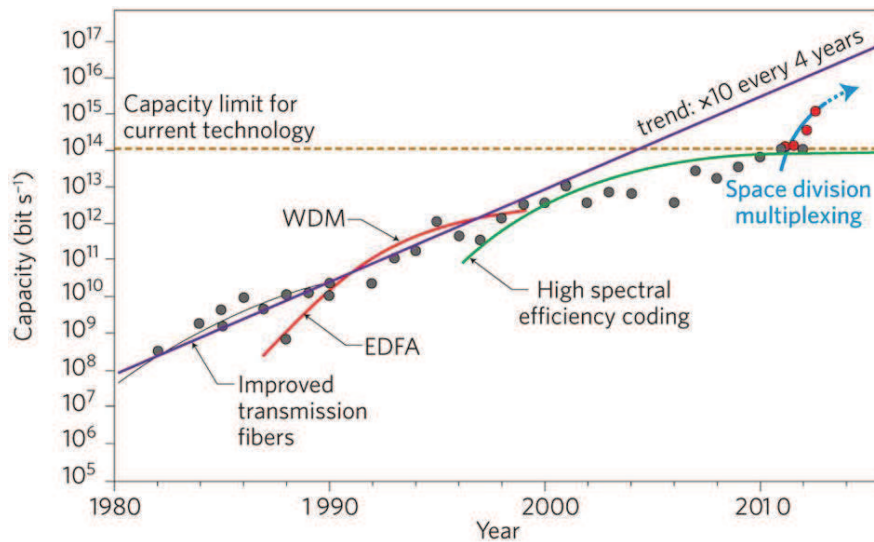


Figure 1.2: The evolution of transmission capacity in optical fibers as evidenced by state of the art laboratory transmission demonstrations over the years. (source: [1]).

In this chapter, we review the evolution of optical transmission systems from the first generations based on simple intensity modulation to the latest promising high-capacity multiple-input-multiple-output (MIMO) solutions. At the same time, we will be presenting the various impairments affecting the performance of each transmission scheme: sources, consequences and the existence or absence of solutions. In section 1.1, we focus on transmission loss, chromatic dispersion and non-linear effects in the fiber as well as on optical noise arising from optical amplification, to which all generations of long-distance optical transmission systems are susceptible. In the following sections, we center on the different multiplexing possibilities WDM, PDM and SDM in an optical fiber and end up with a futuristic look into applications of combined WDM-SDM-PDM transmission systems.

1.1 From on-off keying to multi-level modulations

In its most general form, an optical fiber transmission system is composed of a transmitter, an optical channel containing fiber spans and other optical components, and a receiver. A basic transmitter consists of electrical drivers and optical modulators. Each electrical driver transforms a data bit sequence, that may be coded using a forward error correction (FEC) module, to an electrical signal converted to the optical domain by an electro-optical modulator. The optical signal is transmitted in the fiber link where optical amplifiers are regularly inserted to compensate the fiber attenuation and carry the signal over thousands of kilometers. At the optical receiver, the signal is detected by one or more photodiodes and is converted to the electrical domain. After, analog-to-digital conversion is performed and digital signal processing is applied to compensate the transmission impairments. Finally, the information bits are estimated using an appropriate decoder. A schematic architecture of an optical transmission system is given in Fig. 1.3.

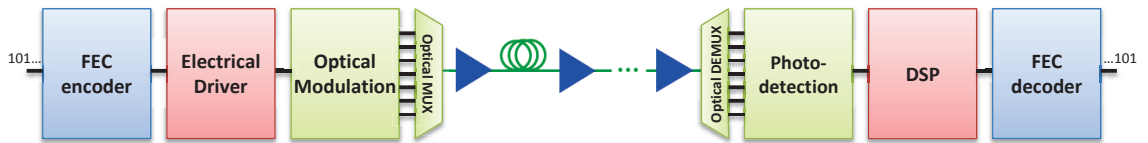


Figure 1.3: The general scheme of a long-distance optical transmission system.

1.1.1 Optical transmitters

A bit sequence at the input of an optical transmitter is converted to an electrical signal that modulates an optical carrier emitted by a laser. The amplitude and phase of the optical signal, at the

output, are determined by a certain number of information bits. The electrical field of the optical signal at the output of a perfect optical modulator is given by:

$$E_0(t) = A_0(t) \exp(\omega_0 t + \phi_0 t) \quad (1.1)$$

where A_0 and ϕ_0 are the amplitude and phase of the output field, $\omega_0 = 2\pi f_0$ is the carrier frequency, in rad/s, linked to the more commonly used wavelength $\lambda_0 = 2\pi c / \omega_0$ with $c = 3 \cdot 10^8$ m/s being the velocity of light in vacuum and f_0 the frequency in Hz. Depending on the used modulation format, we define an amplitude, phase or amplitude and phase modulation.

Real optical modulators are characterized by an electro-optical bandwidth defining the maximum frequency at which they keep responding to the electrical driving signal, an extinction ratio between the maximum and the minimum output power and an insertion loss. Various technologies of optical modulation exist [14] including a direct modulation of the laser or the use of external components that modulate the continuous wave emitted by the laser. These components include electro-absorption modulators, phase modulators and Mach-Zehnder modulators (MZMs). While direct intensity modulation of the laser is the easiest method, it causes an unwanted variation of the lasing frequency, called frequency chirp that induces pulse spreading and hence limits the reach of the transmission system. Electro-absorption modulators, based on the modulation of the absorption coefficient of a semiconductor material, have a reduced frequency chirping effect and can be monolithically integrated with the laser providing a compact and low-cost component called the electro-absorption modulated laser (EML). EMLs are well suited for access networks with typical transmission distances of 80 km and 10 to 40 Gb/s rates. The MZM is another external modulation component that exhibits very low frequency chirping and is a popular device for high-speed long-haul optical transmissions because it offers a large electro-optical bandwidth (up to 40 GHz), small insertion losses (≤ 4 dB) and a high extinction ratio (≥ 20 dB). Moreover, MZMs are known to be weakly wavelength dependent. Their higher cost and power consumption made them unsuitable for access networks.

1.1.1.1 Mach-Zehnder modulator

A MZM is an interferometer composed of two 3 dB couplers and two arms of Lithium Niobate crystal as shown in Fig. 1.4. The modulation occurs by applying a driving voltage to the arms that modifies their refractive index, controlling hence the phase of the light propagating through the arms and creating constructive and destructive interference at the output of the MZM. This interference translates into amplitude fluctuations of the optical signal. The input-output characteristic of an MZM is given by:

$$E_{out} \propto E_{in} \cos\left(\pi \frac{V_1 - V_2}{2V_\pi}\right) \quad (1.2)$$

1. EVOLUTION OF OPTICAL FIBER TRANSMISSION SYSTEMS

where E_{in} and E_{out} are the fields at the input and output of the modulator respectively, V_1 and V_2 are the voltages applied to the arms, and V_π is the drive voltage corresponding to a phase shift from a constructive (maximum) to a destructive (minimum) interference (transmittance). With adequate electrical voltages, MZMs can be used for digital modulations (on-off keying, binary phase shift keying, QAM...) or analog modulations such as OFDM, as shall be seen in the next sections.

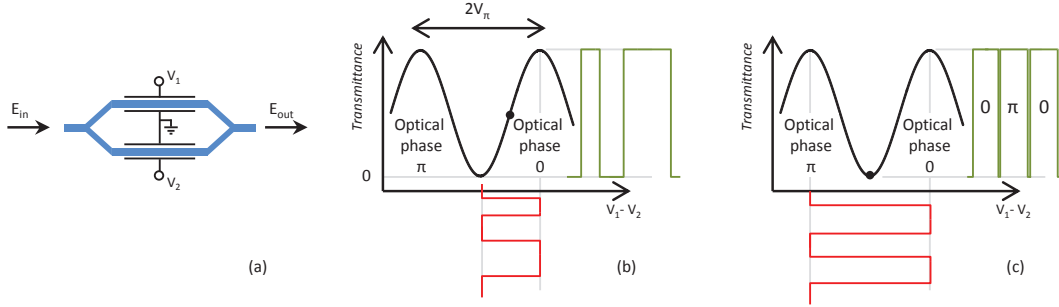


Figure 1.4: Mach-Zehnder modulator: (a) schematic figure of an MZM, (b) Power transmittance with MZM biased at the quadrature point, (c) Power transmittance with MZM biased at the trough point.

1.1.1.2 Modulation formats

A digital modulation format is a set of discernible symbols. Each symbol is defined from a finite set of states. These states can be frequencies and the associated modulation format is known as frequency shift keying (FSK), or amplitudes for amplitude shift keying (ASK), or phases leading to phase shift keying (PSK), or both phases and amplitudes for quadrature amplitude modulation (QAM). In case of amplitude/phase modulation, the combination of states can be represented by a constellation as shown in Fig. 1.5. Each symbol in a q -state constellation can be associated to $k_c = \log_2(q)$ bits.

A modulation scheme is also characterized by the average energy of its constellation denoted as E_S and its bitrate. If each symbol is transmitted during a time T_S , the symbol rate can be defined as $R_S = 1/T_S$ symbols/s (or baud). The corresponding bitrate R_b is related to the symbol rate R_S by: $R_b = \log_2(q)R_S = k_c R_S$.

On-off keying

Non-return-to-zero (NRZ) on-off keying (OOK), consisting in transforming a '0' bit to a weak optical power and a '1' bit to a high power, has been for a long time the preferred modulation format for optical transmission systems. OOK can be achieved using a single MZM driven by an electrical signal with a peak-to-peak amplitude of V_π around a bias voltage set at the quadrature

point as shown in Fig. 1.4-b. At the receiver, a single photodiode followed by a decision threshold that incorporates a clock recovery unit is necessary to retrieve the transmitted bits.

The simplicity and low cost of the NRZ-OOK scheme made it very popular. Nevertheless, this format has a poor noise sensitivity. New formats, with enhanced noise sensitivity and tolerance to non-linearity, were suggested such as the return-to-zero (RZ) pulse-shaped OOK where the power drops from high to weak for the bit '1', or the carrier-suppressed RZ-OOK. The main drawbacks of these methods are their broader spectra compared to NRZ-OOK.

Binary phase shift keying

Phase modulation consists in coding the information on the phase of the signal while maintaining a constant energy profile. In binary phase shift keying (BPSK), a null phase is associated to a '0' and a π -phase to '1'. BPSK has the same spectral efficiency of one bit per information symbol as OOK. Yet, at the same average energy E_S , the distance between the symbols of the BPSK constellation is larger than that of OOK, as can be seen in Fig 1.5, thus providing a 3 dB improved noise sensitivity compared to OOK (noise sensitivity of various modulation formats are presented in appendix A).

A BPSK transmitter is the same as for OOK with a small difference regarding the bias point of the modulator. The MZM is driven by a voltage of $2V_\pi$ peak-to-peak amplitude, and biased at the trough of its characteristic as shown in Fig. 1.4-c. However, the implementation of a BPSK receiver is more complex. At the receiver side, the phase information needs to be optically demodulated before getting to the photodiode that is inherently an intensity detector. Hence, specific receiver schemes are required for phase modulated signals. One way of performing the phase demodulation is differential detection where the phase difference between two consecutive symbols is detected using an interferometer. In this case, the transmitted information is encoded in the phase difference between two successive symbols and the scheme is called differential BPSK (DBPSK). Another way of optically demodulating the differential phase information is using coherent detection that would provide a better sensitivity than a direct detection technique.

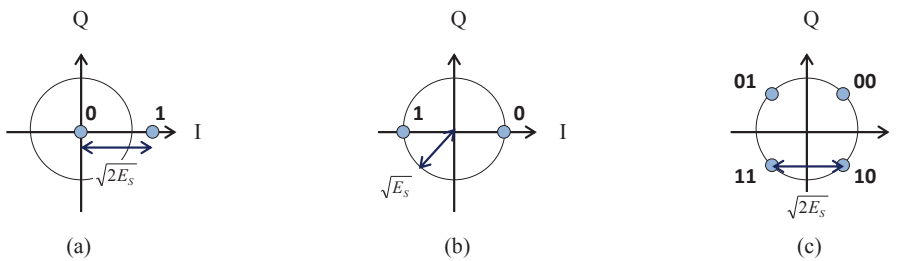


Figure 1.5: Constellations of (a) OOK, (b) BPSK, (c) QPSK or 4-QAM having the same average energy E_S .

Quadrature phase shift keying

The previous modulation formats offer a limited spectral efficiency of a single bit per information symbol. To overcome this limitation, multi-level modulation formats are needed. The first to be investigated in optical communication was quadrature phase shift keying (QPSK) which is a multi-phase modulation format with four distinct symbols $\exp(i\Phi_S)$ having the same amplitude and four equally separated phases as shown in Fig. 1.5-c. Hence, QPSK can encode two bits in each symbol, one in-phase and the other in-quadrature. The bits are mapped to the symbols using a binary Gray code where the bit sequences assigned to two neighboring constellation points differ in only one bit.

For the same target bitrate, QPSK has half the symbol rate of BPSK and hence occupies half its spectral bandwidth, resulting in an enhanced spectral occupancy and making the transmission more tolerant to narrow optical filtering, to chromatic dispersion in the fiber which is quadratically related to the bandwidth as well as to polarization mode dispersion (PMD) as we shall see in the next section. In what concerns noise sensitivity, the minimal distance d_{min} of QPSK, which corresponds to the distance between the closest points in the constellation, is $1/\sqrt{2}$ smaller than that of BPSK at the same average energy E_S . Yet, the spectral efficiency is doubled thus resulting in the same noise sensitivity as BPSK.

A QPSK transmitter can be implemented using two parallel MZMs with a $\pi/2$ phase-shift between their outputs. At the receiver, QPSK also needs an optical demodulation before the photodetection through a differential scheme (DQPSK). Again, coherent detection is another way to optically demodulate DQPSK providing an enhanced sensitivity compared to direct detection.

Quadrature amplitude modulation

Quadrature amplitude modulation (or q -QAM) consists in modulating both phase and amplitude of the signal. The q symbols can be represented by a complex number in polar coordinates: $A_S \exp(i\Phi_S)$. Different flavors of QAM exist such as rectangular QAMs that can be easily implemented with two amplitude modulations on quadrature carriers. The constellation diagrams of a rectangular 4-QAM, that corresponds to a QPSK modulation for the same average symbol energy, 8-QAM, with two equivalent configurations, and 16-QAM modulations are given in Fig. 1.6. The minimal distance of each constellation is also given as a function of the average symbol energy E_S . Although being easy to implement, a rectangular QAM does not maximally space the constellation points at a given energy for $q > 4$, and is thus sub-optimal. Other QAM schemes can be built to optimize the distances between points. For instance, a circular 8-QAM is drawn in Fig. 1.6 showing a larger minimal distance d_{min} than the rectangular 8-QAM at equal average energy. Although optimal non-rectangular QAMs can be found for a given number of points q , they were rarely used

in early generations of optical transceivers because their modulation and demodulation are harder to implement than for rectangular q -QAMs.

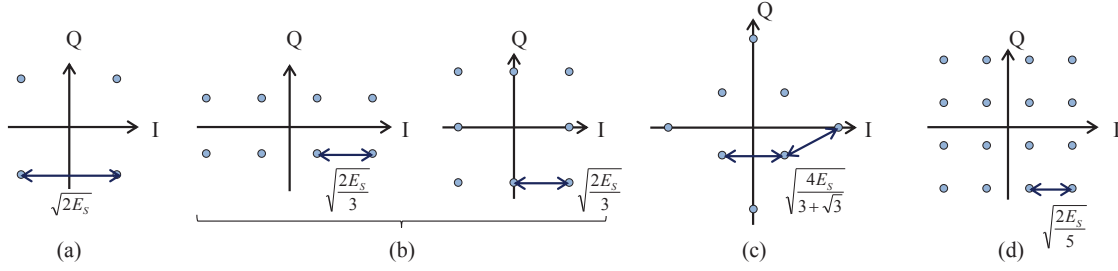


Figure 1.6: Constellation diagrams of: (a) rectangular 4-QAM, (b) rectangular 8-QAM, (c) circular 8-QAM (d) rectangular 16-QAM, with the same average energy E_s .

The spectral efficiency of a q -QAM constellation is given by $k_c = \log_2(q)$. The main advantage of multilevel modulation formats is the increased bitrate over a limited spectral bandwidth. For future optical communication systems, 8-QAM, 16-QAM and up to 128-QAM constellations [15–17] are being considered to achieve a 1.5, 2 and 3.5 capacity increase factor, respectively, compared to QPSK. This capacity increase comes at the cost of complex transmitters and receivers as well as an augmented sensitivity to both linear and non-linear effects. To overcome the complexity at the transmitter, any configuration of multi-level modulation formats (rectangular, circular...) can be implemented using digital-to-analog converters (DACs) instead of multiple MZM branches. Indeed, the availability of high-speed DACs opens the way to a larger set of modulation formats ranging from multi-level modulations to analog modulations such as OFDM.

1.1.2 Propagation in optical fiber transmission systems

A fiber is a cylindrical non-linear waveguide consisting of a core, a cladding and a coating. The light is confined in the core through total internal reflection due to the difference between the refractive indices of the core and the cladding: $n_{cl} \sim 1.444$ for the cladding of silica fibers with a step-index profile and a slightly higher index for the core n_{co} as shown in Fig. 1.7. Depending on the core radius r_c , refraction index difference $\Delta n = (n_{co} - n_{cl})/n_{co}$, and the wavelength λ_0 of the optical carrier, the light waves propagate in distinct patterns also called modes. A mode describes the distribution of the energy of a light wave across the fiber section. A dimensionless parameter V (usually called normalized frequency) defined by the optical and geometrical properties of a fiber determines the number of propagating modes. V is given by [18, Chap.4]:

$$V = \frac{2\pi r_c}{\lambda_0} \sqrt{n_{co}^2 - n_{cl}^2} = \frac{2\pi r_c}{\lambda_0} NA \quad (1.3)$$

1. EVOLUTION OF OPTICAL FIBER TRANSMISSION SYSTEMS

where NA is called the numerical aperture of the fiber. An important parameter for each mode is its cut-off frequency that can be found by solving the propagation equation in a fiber taking into account its optical and geometrical properties as well as the boundary conditions between the core and the cladding. If the cut-off frequency of a mode is smaller than V , the mode is guided by the fiber. In step-index fibers, one mode called the fundamental mode has a zero cut-off frequency and hence is always allowed to propagate. The next mode to appear has a cut-off frequency of 2.405. Hence, the single-mode condition of a step-index fiber is to have $V < 2.405$ [18, Chap.4] and the fiber is called single-mode fiber (SMF). On the other hand, if $V > 2.405$, the higher V goes, the more propagating modes will appear and the fiber is called a multi-mode fiber. In this section, we provide a description of single-mode optical fiber transmission systems with an emphasis on fiber attenuation, chromatic dispersion and the main non-linear effects.

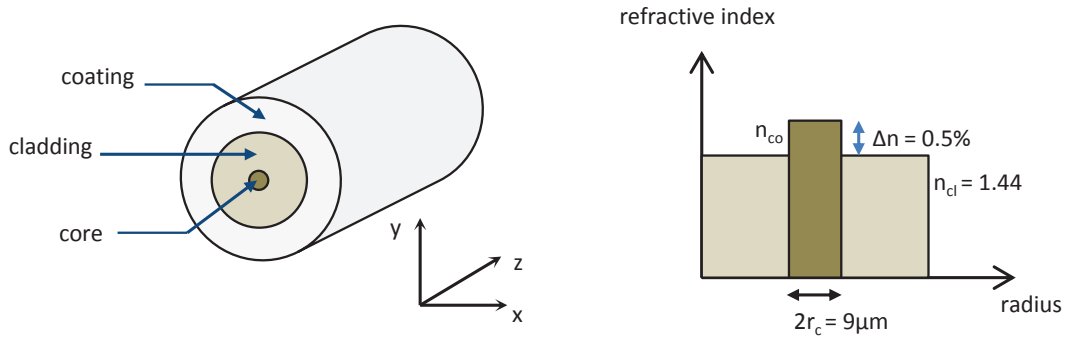


Figure 1.7: Single-mode fiber with a step-index profile.

Propagation in single-mode fibers was thoroughly studied and different models were suggested depending on which effects are intended to be captured. Given a cylindrical core and cladding structure of the fiber with a step-index profile and a small core radius (usually less than $5 \mu\text{m}$ around $1.55 \mu\text{m}$ to satisfy (1.3)), only a single spatial mode can propagate in the fiber at the wavelength λ_0 . The corresponding electrical field E can be expressed in the time domain as [19, Chap.2]:

$$E(x, y, z, t) = F(x, y)A(z, t) \exp(i(\beta z - \omega_0 t)) \quad (1.4)$$

where z is the propagation distance, t the time variable, $F(x, y)$ the normalized spatial distribution of the field in the (x, y) plane and $A(z, t)$ the complex envelope of the field. β is the pulsation-dependent propagation coefficient, approximated by a Taylor series at the angular frequency of the optical carrier ω_0 :

$$\beta_\omega = \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2}\beta_2(\omega - \omega_0)^2 + \dots \quad (1.5)$$

where β_i is the i^{th} derivative of β with respect to ω .

In a standard single-mode fiber, F can be approximated by a Gaussian distribution and $A(z, t)$ contains all the interesting information about the electrical field at a distance z and at a time t . A can be found by solving the non-linear Schrödinger (NLS) equation describing the propagation in the fiber in the time domain [19, Chap.2]:

$$\frac{\partial A}{\partial z} + \beta_1 \frac{\partial A}{\partial t} + \frac{i}{2} \beta_2 \frac{\partial^2 A}{\partial t^2} + \frac{\alpha}{2} A = i \gamma_{NL} |A|^2 A \quad (1.6)$$

where α is the attenuation coefficient of the fiber per unit length and γ_{NL} is the non-linearity coefficient. A major simplification made in (1.6) is the assumption that the polarization state of light does not change during its propagation in the fiber, which is not the case in practice as we shall explain in section 1.2. A more general form of the NLS equation that takes into account the non-linear interactions between two orthogonal polarization tributaries x and y of the optical field is given below where A_x (resp. A_y) is the complex envelope of the electrical field on the x (resp. y) polarization [19, Chap.6]:

$$\begin{aligned} \frac{\partial A_{x/y}}{\partial z} = & \underbrace{-\frac{\alpha}{2} A_{x,y} - \beta_{1x,1y} \frac{\partial A_{x,y}}{\partial t} - \frac{i\beta_2}{2} \frac{\partial^2 A_{x,y}}{\partial t^2}}_{\text{linear effects}} \\ & + \underbrace{i \gamma_{NL} \left(|A_{x,y}|^2 + \frac{2}{3} |A_{y,x}|^2 \right) A_{x,y} + \frac{i\gamma_{NL}}{3} A_{x,y}^* A_{y,x}^2 \exp(\mp 2i\delta\beta_0 z)}_{\text{non-linear effects}} \end{aligned} \quad (1.7)$$

where $\delta\beta_0 = \beta_{0x} - \beta_{0y}$. Due to linear birefringence, group velocities β_{1x} and β_{1y} are different resulting in a polarization dependent dispersion and differential delays between the two multiplexed signals. The $2/3$ factor in the non-linear term comes from the fact that the two considered polarization tributaries are orthogonal. For long-distance transmissions, the last term in (1.7) often changes sign and its contribution averages out to zero and can be neglected.

The conversion between the frequency domain and time domain solutions is done using these equations:

$$A(z, \omega) = \int A(z, t) \exp(-i\Delta\omega t) dt \quad \text{and} \quad A(z, t) = \frac{1}{2\pi} \int A(z, \omega) \exp(\Delta\omega t) d\omega \quad (1.8)$$

We will focus in the following subsections on the different terms in (1.6).

1.1.2.1 Transmission loss

In silica optical fibers, the attenuation of the propagating optical power depends on the wavelength of the signal. The main sources governing the attenuation profile of an optical fiber are Rayleigh scattering and absorption of the material. The former results from microscopic imperfections in

the fiber during its fabrication and affects short wavelengths more than long ones in the near-infrared region (1200 – 1715 μm). The latter is caused by the absorption of silica SiO_2 , strong at long wavelengths as well as by the absorption of OH^{-1} ions at 1.39 μm . These loss sources define two low-loss regions used for optical communications with two minima: a first at 1.55 μm and a second at 1.3 μm . Standard single mode fibers (SSMF), also referred to as G.652, are the most commonly employed for telecommunications and have a low loss coefficient $\alpha_{dB} = 10 \log_{10}(\alpha) = 0.2 \text{ dB/km}$ at 1.55 μm . The optical power is hence divided by 100 after propagating in 100 km of SSMF. For the most recently manufactured fibers, some of the absorption peaks are suppressed and the achieved loss coefficients are as low as 0.16 dB/km.

1.1.2.2 Optical amplification

In long-haul optical transmissions, the signal has to be periodically re-amplified despite the low attenuation of the waveguide. The emergence of optical amplification techniques made this possible without the need to regenerate the optical signal. Erbium doped fiber amplifiers (EDFA), the most popular optical amplifiers, are periodically inserted every 40 to 120km to raise the signal power to the initially injected power at the transmitter, depending on the type of fiber, the total system length and the application (sub-marine or terrestrial systems). They are represented by the blue triangles in Fig. 1.3. Their popularity is mainly due to their large amplification window and relatively flat gain spectrum, thus paving the way for wavelength division multiplexing (WDM) and revolutionizing optical transmission systems.

An EDFA consists of a few-meters long single-mode fiber doped with Erbium Er^{3+} ions in which the optical signal propagates with one or more signals pumped in the core by lasers at 0.98 μm or 1.48 μm corresponding to the absorption band of Erbium. This technology was demonstrated by Desurvire *et al.* [20] and Mears *et al.* [21] in 1987. The pumped power excites the doping ions into a higher energy level, and the ions relax back to the lower level by stimulated emission with the incident optical signal or by spontaneous emission as shown in Fig. 1.8. Stimulated emission leads to the generation of photons with the same properties than the signal photons (same frequency, polarization and phase) which ensures the signal amplification, whereas spontaneous emission leads to the emission of disordered photons in all directions. Some of these photons propagate in the direction of the incident signal and add up as noisy photons, which in turn, excite ions and generate new noisy photons. This process is called amplified spontaneous emission (ASE). Having several optical amplifiers in a long-haul link, the generated ASE noise progressively accumulates, resulting in the principal noise source in these transmission systems.

EDFA gains can reach $\sim 40 \text{ dB}$ with a maximum output power of 23 dBm (0 dBm corresponds to a power of 1 mW). Distinct EDFAs are optimized for the C-band (conventional band: 1.530 – 1.565 μm) and the L-band (long- λ band: 1.565 – 1.625 μm) amplification because the

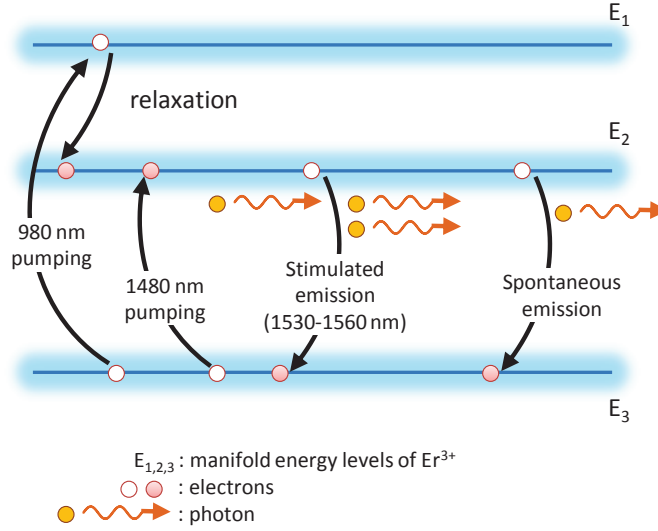


Figure 1.8: Energy level diagram of an EDFA.

latter requires longer Erbium doped fibers. As for ASE noise, it is commonly assumed to be an additive white Gaussian noise. A single EDFA compensating the loss of a fiber of length L kilometers, introduces a noise having the following gain dependent power spectral density for a given polarization mode:

$$N_{ASE} = (G - 1)\bar{h}f_0n_{sp} \quad (1.9)$$

where $\bar{h} = 6.626 \cdot 10^{-34} \text{ J.s}$ is the Planck constant, $G = \exp(\alpha L)$ the gain of the amplifier and $n_{sp} \geq 1$ the spontaneous emission factor that fixes the noise figure NF of an optical amplifier defined as the ratio between OSNRs at its input and output. $NF \approx 10 \log_{10}(2n_{sp})$ at high G and a perfect amplifier with $n_{sp} = 1$ has a noise figure of 3 dB [22, Chap.7]. In practice, EDFAs have a noise figure between 4 and 6 dB.

The total optical noise power, in a link of N_{spans} equally spaced fiber spans with optical amplifiers as in Fig. 1.3, is usually measured within a reference spectral bandwidth B_{ref} of $\Delta\lambda = 0.1 \text{ nm}$ (or equivalently $\Delta f = \frac{c}{\lambda_0^2} \Delta\lambda = 12.5 \text{ GHz}$ at $1.55 \mu\text{m}$) and consequently the optical signal-to-noise ratio (OSNR) is defined as the power ratio between the optical signal and the ASE noise power:

$$\text{OSNR} = \frac{P_S}{2N_{spans}N_{ASE}B_{ref}} \quad (1.10)$$

where P_S is the average optical signal power. The factor 2 corresponds to the ASE power on two orthogonal polarization states.

Gain flattening filters are usually used in optical links to enhance the flatness of the gain spectrum of optical amplifiers. However, EDFAs exhibit a small polarization dependent gain that can

accumulate in long-haul polarization division multiplexed systems and affect their performance, as we shall see later on.

Optical amplification can also be achieved, in a distributed way, using the non-linear stimulated Raman scattering (SRS) effect where the fiber is used as both a transmission and amplification medium [23]. Even though Raman amplification achieves a better effective noise figure and has a wider amplification bandwidth than EDFAs, it also presents a number of disadvantages including a high cost and high-power pumps, hence stricter safety requirements. Hybrid EDFA/Raman amplification schemes were suggested to address these drawbacks, achieving low effective noise figures and extending the maximum reach of the transmission system.

1.1.2.3 Chromatic dispersion

Due to the wavelength dependence of the refractive index of the fiber, the spectral components of the transmitted optical signal travel at different velocities, which is known as material dispersion. Another source of dispersion is waveguide dispersion caused by the propagation of a fraction of the electrical field in the cladding that has different refractive index than the core. Hence, a part of the field travels at a different velocity causing dispersion.

In (1.5), the coefficient β_0 is a constant phase shift and β_1 is the inverse of the group velocity $v_g = 1/\beta_1$ of the electrical field in ps/km, or in other words it is the group delay experienced by a light pulse per unit length. β_2 is related to the group velocity dispersion (GVD) - or chromatic dispersion (CD) - coefficient of the fiber:

$$D = \frac{d\beta_1}{d\lambda} = -\frac{2\pi c}{\lambda_0^2} \beta_2 \quad (1.11)$$

D quantifies the variations of the group delay per unit length with the wavelength. For an SSMF, D is typically 17 ps/(nm.km) at $\lambda_0 = 1.55 \mu\text{m}$. Using (1.11) and considering only chromatic dispersion in (1.6) (ideal loss-free, nonlinear-free fiber) with a change of variable to a reference moving along the fiber at the speed v_g , we can obtain the transfer function of chromatic dispersion in the frequency domain:

$$H_{CD}(z, \omega) = \exp(-i \frac{D\lambda_0^2 z}{4\pi c} \omega^2) \quad (1.12)$$

At a certain wavelength λ_{ZD} , material and waveguide dispersion can cancel each other ($\lambda_{ZD} = 1.32 \mu\text{m}$ for SSMF). Away from λ_{ZD} , dispersion induces pulse broadening that leads to inter-symbol interference (ISI) as shown in Fig. 1.9. This pulse broadening can be evaluated by a dispersion length $L_D = T_0^2/|\beta_2|$ that characterizes the distance after which a Gaussian pulse of initial width T_0 broadens by a factor of $\sqrt{2}$. T_0 being inverse proportional to the symbol rate, it is clear that chromatic dispersion limits the maximum achievable reach of the optical link and can degrade its performance if no CD management techniques are implemented.



Figure 1.9: Inter-symbol-interference due to chromatic dispersion.

CD was first addressed by manufacturing dispersion-shifted fibers with different λ_{ZD} values [19, Chap.1] such that a zero chromatic dispersion is obtained on the transmission wavelength. However, these fibers are more sensitive to non-linear effects. Another way of optically managing CD in long-haul transmission links is the periodical CD compensation through the concatenation of SSMFs and dispersion compensating fibers (DCFs) with a negative dispersion coefficient $D < 0$. However, the non-linear effects are also important within DCF because of their reduced effective core area. Therefore, special dispersion maps with pre-, in-line and post- compensation of CD were studied to enhance the tolerance to non-linear effects [24]. Later on, with the return of coherent detection, digital equalization techniques were applied to completely compensate CD which is, as can be seen from its transfer function in (1.12), a “deterministic” and unitary effect [25]. In fact, if we further develop (1.5), we obtain β_3 defined as the GVD slope or the variation of GVD as a function of wavelength which leads to a quasi-deterministic CD.

1.1.2.4 Non-linear effects

The non-linear coefficient γ_{NL} in (1.6) is given by:

$$\gamma_{NL} = a \frac{n_2 \omega_0}{c A_{eff}} \quad (1.13)$$

where $a = 8/9$ for SSMF and $a = 1$ for polarization maintaining fibers [26], A_{eff} is the effective core area over which the signal power is assumed to be constant and n_2 is the non-linear refractive index. Indeed, the refractive index of silica changes with the intensity of the propagating signal, a phenomenon known as the optical Kerr effect. n_2 has typical values of $2 - 3 \cdot 10^{-20} \text{ m}^2 \text{W}^{-1}$ and A_{eff} varies between $20 - 50 \text{ } \mu\text{m}^2$ for DCFs and $80 - 100 \text{ } \mu\text{m}^2$ for SSMFs around $1.5 \text{ } \mu\text{m}$ depending on the fiber design. Having seen that the fiber is a lossy medium, we expect that non-linear effects will be more important in the first kilometers of a span where the optical power is still high. The effective fiber length where non-linear effects are strong is denoted L_{eff} and approximated by [22, Chap.7]:

$$L_{eff} = \frac{1 - \exp(-\alpha L)}{\alpha} \quad (1.14)$$

where L is the total length of the fiber span.

1. EVOLUTION OF OPTICAL FIBER TRANSMISSION SYSTEMS

We can see from (1.6) that the term containing γ_{NL} will induce a non-linear intensity-dependent phase shift in the electrical field A . This effect is known as self-phase modulation (SPM). Since the signal intensity can vary with time (OOK, transitions from a symbol to another in BPSK, QPSK), the non-linear phase shift is also time-varying, and thus generates a time-varying frequency shift referred to as frequency chirping. By shifting the signal frequencies towards lower and upper values, SPM causes spectrum broadening. The shifted frequencies will propagate at a different speed due to chromatic dispersion, interfere with other signal components and induce non-linear amplitude distortions.

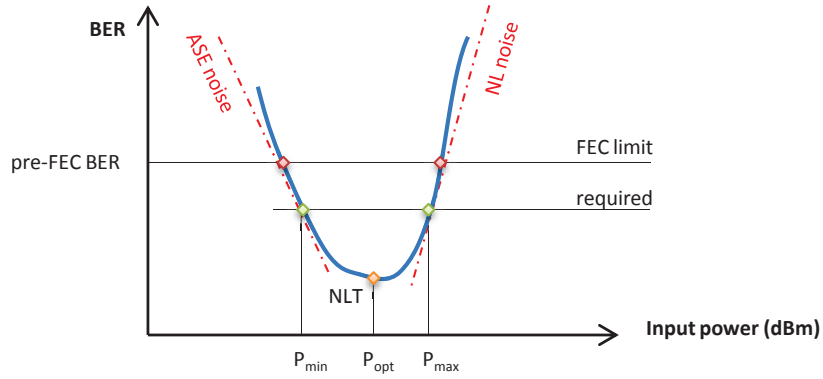


Figure 1.10: Characteristic bath-tub shaped BER performance as a function of the injected input power per wavelength in the optical fiber link. Extra margins $[P_{min} : P_{max}]$ may be considered rather than a single operating point P_{opt} to take into account system aging.

To sum up, optical transmission systems are limited by optical amplification noise for low OSNR values, but also by non-linear intensity-dependent effects that become dominant with an increasing input power as can be seen in Fig. 1.10. Hence, a trade-off between these two limits must be found and the system must be operated at an optimal input power. This optimal operation point is referred to as the non-linear threshold (NLT). The NLT of an optical transmission system of total length L_{tot} is simulated or measured by searching for the input power that ensures the best bit error rate (BER) performance. It varies with the modulation format, the bandwidth, the dispersion map and the optical components in the link.

1.1.3 Optical receivers

1.1.3.1 Direct detection schemes

Direct detection (DD) is commonly used in 10 Gb/s and 40 Gb/s OOK and DPSK optical transmission systems. For OOK signals, one photodiode acting as a quadratic receiver provides an electrical current proportional to the intensity of the received optical signal. Then, a clock recovery circuit provides the right samples to a threshold decision element. For PSK modulations, in

order to keep the receiver complexity low, differential schemes were used as mentioned in section 1.1.1.2 along with single-symbol-delay optical interferometers at the receiver to perform the optical demodulation before photo-detection. Balanced photodiodes are connected to the outputs of the interferometers providing an electrical current that is proportional to the phase shift between two successive symbols. The different direct detection schemes can be seen in Fig. 1.11.

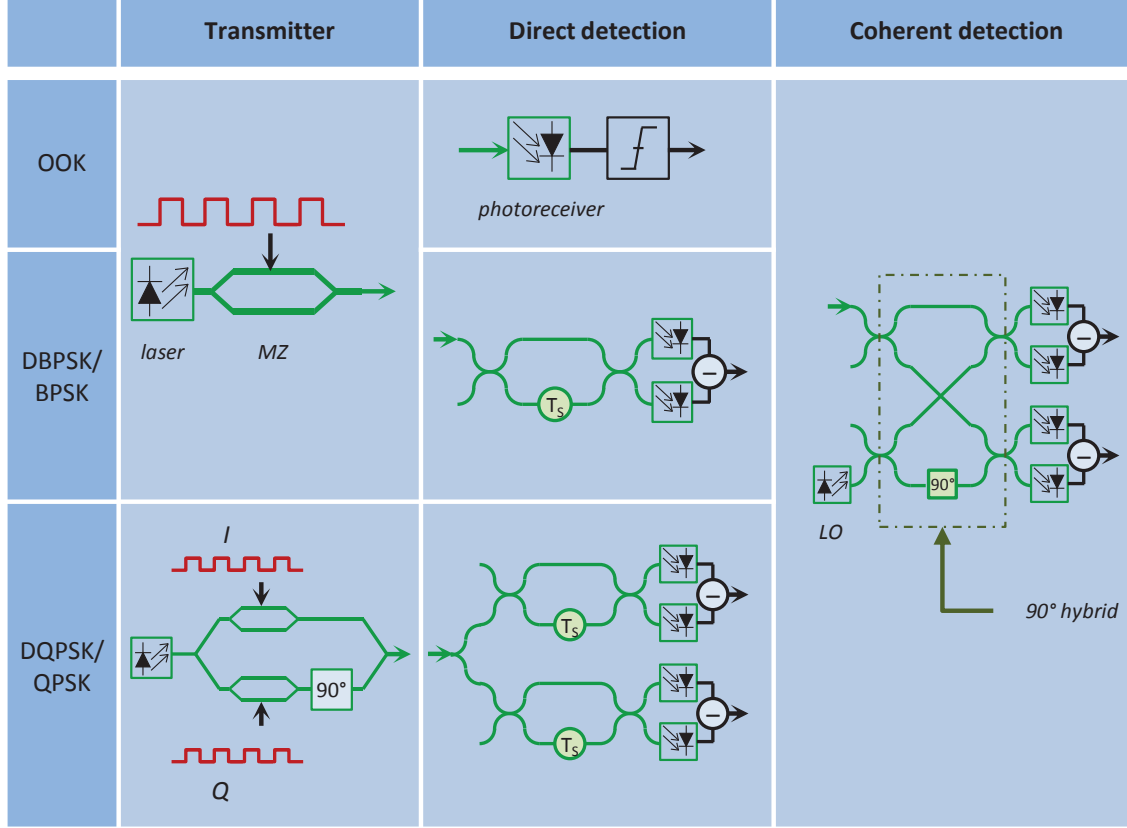


Figure 1.11: Transmitters, direct detection and coherent detection receivers for various modulation formats.

Despite its simple architecture, the performance of direct detection for differential encoded BPSK and QPSK can be severely affected by receiver imperfections such as temporal or gain imbalances between the different arms of the interferometers [27]. Besides, the receiver architecture will become complex for higher spectral-efficiency modulations. Above that, DD systems are limited by linear polarization mode dispersion effects that cannot be efficiently mitigated after photo-detection due to the absence of information on the phase and polarization of the optical signal.

1.1.3.2 Coherent detection schemes

Coherent detection consists in beating the incident optical signal with a local oscillator laser (LO) oscillating at nearly the same frequency and serving as a reference signal which provides a better

noise sensitivity than direct detection [28] and differential detection [29]. Coherent detection was first investigated in the 1980s [30]. Yet, the invention of EDFAs on one hand and the need for optical phase locking required to compensate the phase shift between the LO and the optical carrier on the other hand, made its implementation complex and unnecessary. However, later, the interest in coherent detection was revived by the advancements in the increased performance, reliability and speed of dedicated integrated circuits that permitted to replace optical phase-lock loops with digital carrier recovery techniques.

In order to recover the in-phase and in-quadrature components of the signal, a 90°-hybrid phase diversity intradyne coherent receiver is usually used where the local oscillator frequency ω_{LO} is close to the carrier frequency ω_0 . The LO is separated into two 90°-phase shifted parts that interfere with the incident optical signal. Then, balanced photodiodes provide two currents corresponding to the I and Q components of the received signals [28]:

$$\begin{aligned} I_Q(t) &= \mathfrak{r} \sqrt{P_0(t)P_{LO}(t)} \cos(\Phi_0(t) - \Phi_{LO}(t)) \\ I_I(t) &= \mathfrak{r} \sqrt{P_0(t)P_{LO}(t)} \sin(\Phi_0(t) - \Phi_{LO}(t)) \end{aligned} \quad (1.15)$$

where $P_0(t), P_{LO}(t)$, Φ_0 , Φ_{LO} are respectively the powers and phases of the received signal and the local oscillator, and $\mathfrak{r}$ is the responsivity of the photodetection. In (1.15), we suppose that the carrier offset $\Delta\omega_{IF} = \omega_0 - \omega_{LO}$ is null or perfectly compensated (homodyne detection). The scheme of a coherent receiver can be seen in Fig. 1.11. The amplitude, phase (and polarization as we shall see later in this chapter) of the received signal being now available in the electrical domain, coherent detection opens the way to the digital mitigation of the various linear and non-linear impairments of the optical transmission link, and in addition, to the use of new modulation formats such as higher order modulations and orthogonal frequency division multiplexing. Hence, the available bandwidth can be better utilized than with binary modulation formats and the capacity of the transmission system can be boosted without changing the installed fibers.

1.1.3.3 Digital signal processing

Digital signal processing can be applied either at the transmitter to pre-distort the optical signal or at the receiver to equalize linear and/or non-linear impairments. In direct detection schemes, pre-compensation of chromatic dispersion was demonstrated [31] but these systems were more vulnerable to non-linear effects. CD compensation at the receiver was also proposed [32], still, the absence of phase information and the non-linear quadratic detection limited the performance of DD systems. With the introduction of coherent detection, all the information on the complex amplitude of the optical signal is linearly transposed to the electrical domain. Therefore, optical filtering, distortion compensation and more sophisticated digital processing can now be performed digitally at the receiver.

The DSP circuit of a single-carrier digital coherent receiver must perform a sequence of important operations to retrieve the transmitted information. First, the currents in (1.15) must be sampled at twice the maximum bandwidth in order to satisfy the Shannon-Nyquist theorem, and quantized by analog-to-digital converters (ADCs). Next, clock recovery and re-timing of the signals are performed before equalizing the impairments of the optical channel [25]. The equalization step contains a chromatic dispersion compensation module. Then, the carrier phase noise is removed using the Viterbi and Viterbi algorithm [33] when a constant-modulus modulation is used such as QPSK, and finally the symbols are demodulated and decoded. A general scheme of a single-carrier digital coherent receiver is given in Fig. 1.12.

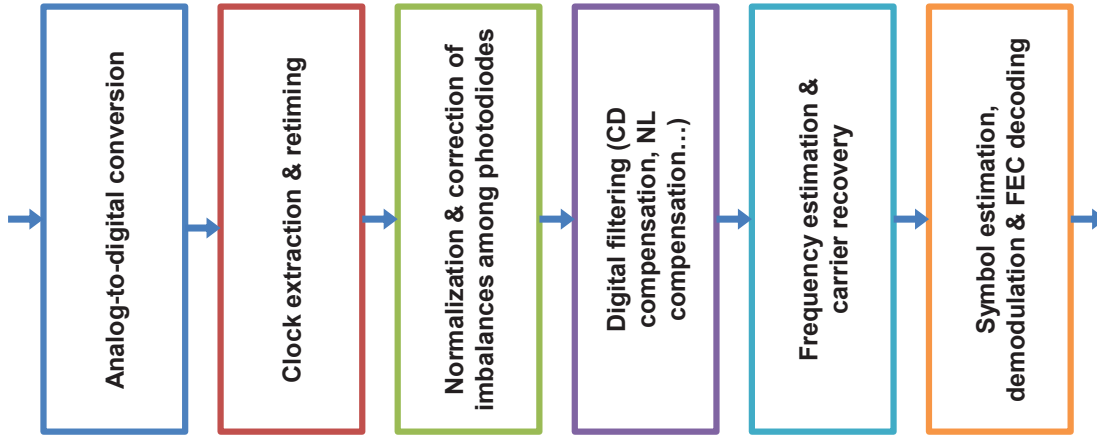


Figure 1.12: Schematic of the DSP blocks in a single-carrier digital coherent receiver.

Whether CD compensation is realized in-line using DCFs or not, all or a residual CD can be compensated at the receiver side using a temporal finite impulse response (FIR) filter [25] or a single-tap filter applied on the frequency domain signal obtained through a Fast Fourier Transform (FFT) module. In the former case, the FIR filter must have a number of taps at least as long as the maximum delay spread caused by CD. In the latter case, the dispersion compensating filter is the inverse of the CD transfer function in (1.12) and an arbitrary amount of CD can be compensated. The quasi-deterministic nature of CD makes its complete digital equalization with a static filter very interesting. Indeed, this will keep us from using DCFs with a negative dispersion parameter that usually have large non-linearity coefficients. For long-haul transmissions, a large number of taps will be required for full temporal CD compensation at the receiver while the complexity of the single-tap frequency domain filter will be limited to the implementation of the FFT and inverse FFT (iFFT) blocks.

Nevertheless, the described coherent receiver is a polarization insensitive detection technique and requires active polarization control in order to properly operate. We shall see later that a polarization-diversity version of coherent detection along with appropriate DSP will solve this issue and give access to an additional degree of freedom in SSMF which is polarization, thus

allowing to double the spectral efficiency of the installed links and to compensate polarization dependent impairments.

1.2 Wavelength Division Multiplexing (WDM)

Empowered by the manufacture of powerful wide-band optical amplifiers, WDM consists in transmitting independent data streams using different wavelengths, also called “channels” that propagate simultaneously in the same fiber as shown in Fig. 1.13. WDM expanded the system capacity a hundred-fold and supported the traffic growth with a very modest cost per bit thanks to the large available THz bandwidth. A major advantage of WDM is the absence of linear crosstalk between channels even in dense WDM systems with 50 GHz and 25 GHz wide channels according to the ITU-T G.694.1 recommendation. The channels can be individually filtered before being combined at the transmitter and before being detected at the receiver. Using both the C- and L- bands, up to 220 channels can be transmitted on a 50 GHz grid.

However, WDM systems are subject to non-linear inter-channel crosstalk induced by cross-phase modulation (XPM) and four-wave mixing (FWM). The former arises from optical Kerr effect similarly to SPM, with intensity variations of a wavelength ω_1 altering the phase of another channel ω_2 . The latter occurs when two co-propagating wavelengths ω_1 and ω_2 create two new components at $\omega_3 = 2\omega_1 - \omega_2$ and $\omega_4 = 2\omega_2 - \omega_1$ that may coincide with other data-carrying wavelengths, and hence cause interference and information loss [19, Chap.10].

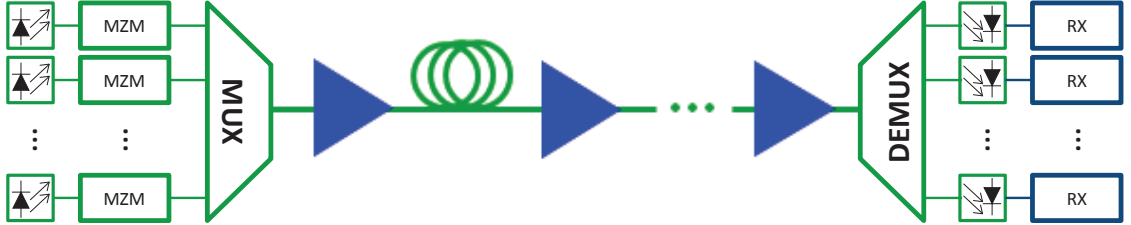


Figure 1.13: Wavelength division multiplexing transmission scheme.

XPM-induced penalties can be reduced by increasing the channel separation in WDM systems. Unequal channel spacing [34] and special dispersion maps with local non-zero dispersion and an average overall null chromatic dispersion [35] are effective techniques to combat FWM. Yet, with an increasing number of channels, every single channel is disturbed by many others and the induced penalties will depend on many parameters, for instance the channel spacing, the used modulation formats, the dispersion coefficients and the channel bitrate. Indeed, for low bit rates and/or low chromatic dispersion, the signals interact during a significant time inducing important non-linear phase shifts. The higher the bit rates and/or the chromatic dispersion, the shorter this

interaction becomes. However, the bandwidth of each channel is wider for a high bitrate and non-linear effects similar to XPM and FWM will arise inside each channel [36]. These effects are referred to as intra-channel XPM and FWM (IXPM and IFWM).

The mitigation of these non-linear techniques using powerful DSP capacities and coherent detection is currently investigated. Back-propagation and Volterra-based techniques trying to invert the propagation equation (1.6) are among the most powerful, yet complex, solutions [37].

1.3 Polarization Division Multiplexing (PDM)

So far, the amplitude, phase and wavelength dimensions were employed to transmit information. The use of a new dimension “polarization” will enable us to reach a 100 Gb/s bitrate per wavelength by sending QPSK data symbols over two orthogonal polarization states, hence reducing the required bandwidth by a factor of two when compared with the previous single-polarization systems [38]. This new solution is referred to as polarization division multiplexing (PDM) or polarization multiplexing (Pol-Mux) and has established itself as the most widely used polarization-sensitive scheme. Apart from the more efficient utilization of the available bandwidth, coherent PDM solutions have increased tolerances to linear and nonlinear transmission effects than the previous schemes and inherit all the benefits of coherent detection.

Polarization is a property of electromagnetic waves that describes the direction of the transverse electric field. This field can be divided into two orthogonal components x and y in an arbitrary frame of reference. The state of polarization (SOP) is defined as the pattern traced out by the electrical field in the transverse plane as a function of time. In an optical fiber, the two orthogonal components can be interpreted as the fiber supporting two orthogonal channels. The SOP is then defined by the relative amplitude and phase of the complex envelope of the orthogonal polarization components x and y given by:

$$\mathbf{a} = \begin{pmatrix} A_x \\ A_y \end{pmatrix} = \begin{pmatrix} \sqrt{P_x} \exp(i\phi_x) \\ \sqrt{P_y} \exp(i\phi_y) \end{pmatrix} \quad (1.16)$$

also known as the Jones vector. P_x , P_y , ϕ_x , ϕ_y are respectively the powers and phases of the orthogonal components A_x and A_y . More on the polarization of light can be found in appendix B.

It is important to note that PDM is only one flavor of a larger group of polarization sensitive modulation formats that includes bit-wise alternating-polarization (APol) modulation or polarization-switched QPSK (PS-QPSK) [39] and polarization shift keying (PolSK). Yet, PDM was chosen for optical transmission systems because it has a spectral efficiency advantage over APol where some bits are needed to indicate a polarization switch hence introducing redundancy. PDM has also a lower-complexity TX/RX structure compared to PolSK that needs a precise control of the SOP of the electrical field defined in the 3-dimensional Stokes space to encode the data

1. EVOLUTION OF OPTICAL FIBER TRANSMISSION SYSTEMS

and triple the spectral efficiency compared to a single-polarization scheme. For the definition of SOP in the 3-D Stokes space, please refer to appendix B.

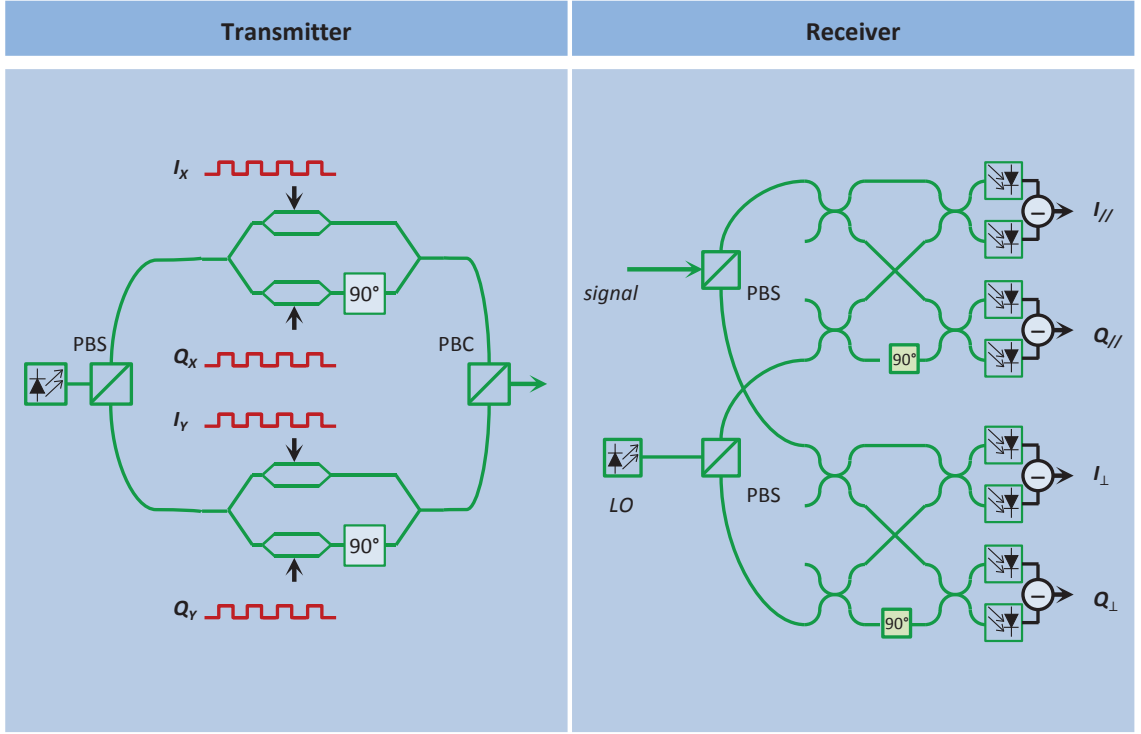


Figure 1.14: Architecture of dual polarization transmitter and corresponding coherent receiver.

1.3.1 Principle of coherent PDM transmissions

PDM is realized by sending independent data in each of two orthogonal linear polarization states in the optical fiber, thus doubling the spectral efficiency of single-polarization transmissions. At the receiver, the data streams can be separated using a polarization sensitive direct detection receiver [40], however a dual-polarization coherent receiver shown in Fig. 1.14 will give access to the full information on the optical field (amplitude, phase, polarization state) and enable the use of better-performing digital equalization of the channel impairments [25, 41–43]. A general transmission scheme of a coherent PDM-QPSK system with a spectral efficiency of 4 bits per time slot can be seen in Fig. 1.15. A close-up on the dual polarization coherent receiver was previously given in Fig. 1.14.

High bitrate PDM optical transmissions suffer several linear and non-linear impairments that include - in addition to chromatic dispersion, fiber attenuation and non-linear SPM - random birefringence and polarization mode dispersion (PMD) causing a constant mix of the data-carrying polarizations and inducing differential delays between the polarization tributaries, random differential losses known as polarization dependent Loss (PDL) and non-linear cross phase modulation

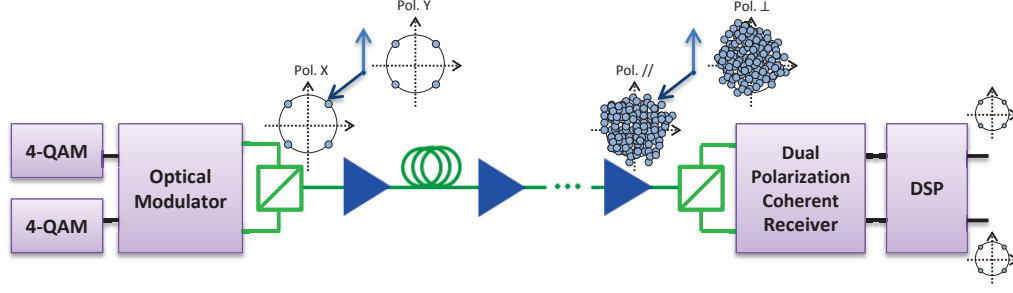


Figure 1.15: General scheme of a Pol-Mux optical transmission.

(XPM).

The linear effects can be modeled using a transfer matrix approach. We commonly refer to the transfer matrix of an optical component $\mathbf{H}(\omega)$ as the Jones matrix which is a 2×2 complex matrix representing the effect of the optical component on the electrical field in the frequency domain [44]:

$$\begin{bmatrix} A_X^{out} \\ A_Y^{out} \end{bmatrix} = \mathbf{H}(\omega) \begin{bmatrix} A_X^{in} \\ A_Y^{in} \end{bmatrix} \quad (1.17)$$

Any optical link - in the linear domain - can be expressed as a concatenation of frequency-domain Jones matrices which are of two kinds: unitary matrices (for dispersive effects) and positive Hermitian matrices (attenuation, PDL). This approach turns a polarization-multiplexed optical channel into a 2×2 multiple-input-multiple output (MIMO) channel, thus providing the ground for new equalization schemes. Non-linear polarization dependent effects also arise when the power of the injected signals becomes important. These effects are modeled in the time domain. Consequently, the propagation in the optical channel can be modeled using the split-step Fourier method that will be explained and used in chapter 3. In the following, we look at the sources, modeling and management of the linear and non-linear effects that impair a long-haul PDM optical transmission system.

1.3.2 Propagation equations

The propagation equations that govern the evolution of two orthogonal polarization tributaries along a fiber were given in (1.7). While the two polarizations have the same β_2 and γ_{NL} , they experience two different group velocities β_{1x} and β_{1y} due to linear birefringence of the optical fiber. SSFMs used in most of optical transmission systems exhibit a small birefringence that randomly scatters the polarization state of the optical field over long distances. This is not the case for all types of fibers such as short segments of polarization maintaining fibers used in some optical components. Hence, the resulting non-linear effects can be modeled by averaging (1.7) over all

possible polarization changes that leads to the Manakov equation given by [19, Chap.6,10]:

$$\frac{\partial A_{x/y}}{\partial z} = -\frac{\alpha}{2}A_{x,y} - \beta_{1x,1y}\frac{\partial A_{x,y}}{\partial t} - \frac{i\beta_2}{2}\frac{\partial^2 A_{x,y}}{\partial t^2} + i\frac{8}{9}\gamma_{NL}(|A_{x,y}|^2 + |A_{y,x}|^2)A_{x,y} \quad (1.18)$$

Other linear interactions between the two polarizations tributaries that occur when the signal propagates in the optical components of the link are not depicted in this equation but will be considered in the next sections.

1.3.3 Polarization Mode Dispersion (PMD)

1.3.3.1 PMD sources and effects

In an ideal fiber with no imperfections and a perfectly circular symmetry, both orthogonal polarizations have an identical group delay and the signal on one polarization will not couple into the orthogonal polarization. However, in a real SSF fiber, slight geometrical imperfections such as asymmetries in the manufacturing process as well as mechanical and thermal stresses make the fiber birefringent, meaning that the refractive index will depend on the polarization and propagation direction of the incident light, resulting in a difference in the group velocities $\Delta\beta_1 = |\beta_{1,x} - \beta_{1,y}|$ between the two orthogonal polarizations.

These imperfections cause a random spreading and overlap of the information-carrying polarizations and lead to polarization dependent delays, pulse broadening and inter-symbol interference (ISI) in the fiber. While geometrical imperfections may be constant, thermal and mechanical stresses vary over time and along the fiber, hence the accumulated differential group delay (DGD) is not constant. This phenomenon is referred to as polarization mode dispersion (PMD). Field measurements and studies of the dynamics of PMD in optical fiber links revealed that the accumulated DGD varies on a millisecond scale for fibers with strong PMD [45] to the scale of a second in installed fibers [46].

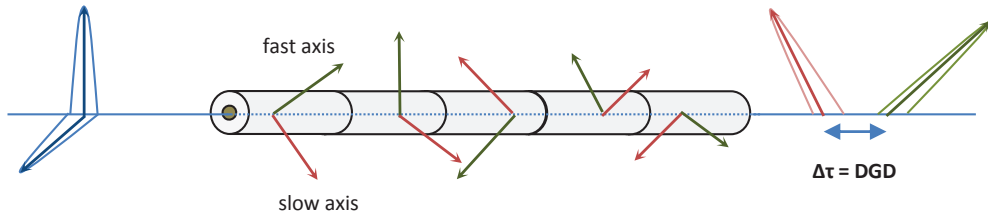


Figure 1.16: Modeling PMD in the fiber. The incident polarization tributaries are rotated and delayed by the birefringent elements, but remain orthogonal at the output.

A fiber span can be modeled as a composite Jones matrix made of a concatenation of independent randomly oriented birefringent components as shown in Fig. 1.16 and given by:

$$\mathbf{F} = \prod_{i=1}^N \mathbf{R}_i \mathbf{D}_i = \prod_{i=1}^N \begin{bmatrix} \cos(\alpha_i) & \sin(\alpha_i) \\ -\sin(\alpha_i) & \cos(\alpha_i) \end{bmatrix} \begin{bmatrix} \exp(-i(\omega\tau_i + \phi_i)/2) & 0 \\ 0 & \exp(+i(\omega\tau_i + \phi_i)/2) \end{bmatrix} \quad (1.19)$$

Each birefringent component $\mathbf{D}_i$ works as a phase shifter and will change the polarization state of the incoming signal on the Poincaré sphere. The phase shift is constant in each element and is induced by two effects: $\phi_i \in [0 : 2\pi]$ caused by the local birefringence of each optical component (phase birefringence) and $\omega\tau_i/2$ a frequency-dependent phase shift resulting from PMD, τ_i being the differential group delay. The rotation matrix $\mathbf{R}_i$ models a random mismatch of $\alpha_i \in [0 : 2\pi]$ between the incident polarization states of the signal and the principal polarization states (PPS) of the birefringent element.

The total PMD-induced pulse broadening depends on a correlation length l_c over which two polarization components remain correlated [19, Chap.6]. For short distances where $z \ll l_c$, the mean accumulated DGD $\overline{\Delta\tau}$ scales linearly with z as is the case of polarization-maintaining fiber. For multi-kilometer distances $z \gg l_c$, due to the random imperfections along the SSF, PMD accumulation corresponds to a random walk [46] and the mean accumulated DGD $\overline{\Delta\tau}$ is proportional to the square root of the propagation distance L in kilometers:

$$\overline{\Delta\tau} = PMD\sqrt{L} \quad (1.20)$$

where PMD is the PMD coefficient of a fiber measured in $\text{ps}/\sqrt{\text{km}}$ that assesses the strength of polarization dispersion. PMD coefficients as low as $0.05 \text{ ps}/\sqrt{\text{km}}$ can be achieved [46] compared to $1 \text{ ps}/\sqrt{\text{km}}$ in fibers from 1980s. However, DGD values larger than the average can occur (when there are points along the fiber with a high local birefringence due to tight bends or when some elements in the link exhibit polarization dependent loss) and the PMD must be addressed to avoid ISI. Indeed, the impact of PMD on the system performance increases linearly with the symbol rate. A 100 Gb/s system is therefore 10 times more vulnerable to PMD than a 10 Gb/s system.

1.3.3.2 PMD management

Initially, optical mitigation methods were proposed for PMD. The solutions consisted in using dynamically adjusted birefringent elements to compensate PMD and align the polarizations to the frame of reference of the optical receiver. Yet, these methods are costly, are unable to track fast polarization changes and require a precise active control of the optical compensators [47]. Later, with the first applications of DSP algorithms in DD systems, maximum likelihood sequence estimators (MLSE) based on the Viterbi algorithm [48] showed a good performance against both chromatic dispersion and PMD [32]. In short, the algorithm searches among all possibly emitted

sequences, the one that maximizes the probability to obtain the received sequence. This algorithm requires an initial training sequence to estimate *a priori* probabilities, i.e. the probabilities of emitting a certain sequence of bits and receiving its changed version.

Having access to the full information on the received electrical field with coherent detection, a larger choice of DSP algorithms is available. MIMO processing including polarization demultiplexing and interference-eliminating techniques were suggested and have proven their ability to completely mitigate PMD and offer robust transmission systems. These techniques include single-carrier solutions with adaptive time-domain equalization [25] and multi-carrier techniques such as orthogonal frequency division multiplexing (OFDM) [43]. We will explain and compare both of these techniques in the next chapter.

In 2010, Sano et al. [49] achieved a record transmission capacity of 69 Tb/s (432 wavelengths and 171 Gb/s per wavelength) over 240 km of optical fibers thanks to 16-QAM PDM and coherent detection followed by MIMO DSP. More recently, in 2014, a 256 Gb/s per wavelength rate was achieved over 960 km [15] also using 16-QAM PDM and digital coherent detection. Dispersive effects have been extensively studied in the last years whereas the mitigation of another far more detrimental effect: polarization dependent loss was still an open issue.

1.3.4 Polarization Dependent Loss (PDL)

1.3.4.1 PDL sources and effects

PDL (or PDG for polarization dependent gain) refers to the polarization dependence of the insertion loss (or gain) of an optical element. It is mainly introduced by non-ideal in-line optical components spread along the link such as amplifiers, isolators, couplers and re-configurable optical add-drop multiplexers (ROADMs) rather than the fiber itself. The transfer matrix of a PDL element shown in Fig. 1.17 is given by:

$$\begin{aligned} \mathbf{H}_{PDL} &= \mathbf{R}_\alpha \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\epsilon} \end{bmatrix} \mathbf{R}_\alpha^{-1} \\ &= \frac{1}{\sqrt{1+\gamma}} \mathbf{R}_\alpha \begin{bmatrix} \sqrt{1+\gamma} & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \mathbf{R}_\alpha^{-1} \end{aligned} \quad (1.21)$$

The diagonal matrix gives the imbalanced attenuation values of the least and most attenuated polarization states corresponding to the lossless and the lossy axis respectively. The rotation matrix $\mathbf{R}_\alpha$ describes a random mismatch of axes between the polarization states of the incident signal and the principal polarization states (PPS) of the PDL component. ϵ and γ are defined through:

$$\Gamma_{dB} = 10 \log_{10} \frac{1+\gamma}{1-\gamma} = -10 \log_{10}(\epsilon) \quad (1.22)$$

where $\Gamma_{\text{dB}} \geq 0$ is the PDL coefficient in dB (simply referred to as PDL) and consists of the ratio between the highest and the lowest coefficients of the diagonal matrix of $\mathbf{H}_{\text{PDL}}$. The normalization factor in the model using the variable γ is commonly dropped in literature [50, 51] because the diagonal matrix and the rotation matrices are enough to take into account the power imbalance between the polarization states as well as their crosstalk.

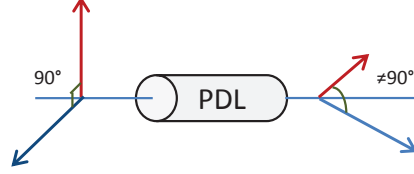


Figure 1.17: PDL element inducing unequal losses for the polarization tributaries of the input signal and breaking their orthogonality.

One major source of PDL is the EDFA where the gain is dependent on the alignment of the polarizations of the pump lasers and the signal. This is called polarization dependent gain (PDG). Fortunately, the birefringence in the optical fiber randomly varies the polarization state of the signal along its length, which tends to average out PDG to typically ≤ 0.5 dB. This residual PDG is also induced by optical isolators in the EDFAs. Hence, PDG is very difficult to observe in a single amplifier, but is noticeable in long-haul transmissions with several cascaded amplifiers.

What we have mentioned for EDFAs is also valid for other in-line optical components exhibiting PDL. The individual PDL value γ of each component is kept as low as possible, i.e. $0 \leq \gamma \ll 1$. However many optical components with PDL are found in long-haul optical links and can lead to large accumulated PDL values (up to 3 dB measured in an installed link [52]). These components connect fibers that act mainly as birefringent elements on the propagating polarization-multiplexed signal, inducing random polarization couplings. The combination of distributed PDL and birefringence leads to a random overall PDL value that needs to be statistically described. It prevents, at the same time, a rapid linear increase of the accumulated PDL with the number of PDL elements in the link. Indeed, PDL statistics were studied in previous works, especially by A. Mecozzi, M. Shtaif [53] and N. Gisin [54]. In [53], it is found that the probability density function of Γ_{dB} is a Maxwellian distribution when we consider many identical low-PDL components distributed in the optical link such that:

$$\text{pdf}(\Gamma) = \sqrt{\frac{2}{\pi}} \frac{\Gamma^2 \exp(-\Gamma^2/(2a^2))}{a^3} \quad (1.23)$$

defined over $\Gamma = [0 : +\infty[$ with a positive parameter $a > 0$. The mean of this distribution is given by $\mu = 2a\sqrt{2/\pi}$ and variance $\sigma^2 = a^2(3\pi - 8)/\pi$. The Maxwellian distribution approximation is valid as long as $\Gamma \ll 34\text{dB}$ and a is linked to the number of PDL components and their individual PDL value.

Imbalanced power & loss of orthogonality

Unlike the unitary dispersive effects (CD, PMD) that cause no loss of energy to the transmitted signal, PDL causes OSNR distortions. The information-carrying polarizations experience unequal loss and crosstalk [50]. From (1.21), we can see that the loss affecting each channel depends on the incident polarization states. Two particular cases are noteworthy:

- If $\alpha \equiv \{0, \pi/2\}(\text{mod } \pi)$, one polarization state will endure the maximum loss while the other will be privileged, resulting in a maximum gain offset between the two polarizations that remain orthogonal as shown in the left part of Fig. 1.18.
- If $\alpha \equiv \{\pi/4, 3\pi/4\}(\text{mod } \pi)$, the incident polarization states will observe the same loss. However, their orthogonality defect would be maximal as shown in the right part of Fig. 1.18.

For intermediate values, the polarization states will suffer a mix of these two impairments: gain offset and loss of orthogonality [55]. These impairments result in an OSNR penalty and a bit error rate (BER) degradation for a digital coherent receiver. A constant Γ_{dB} of 3 dB induces a 2 dB OSNR penalty at a BER of 10^{-3} [50]. Penalties larger than 1dB were also experimentally observed in a $10 \times 100\text{km}$ link with accumulated PDLs greater than 2 dB in [56].

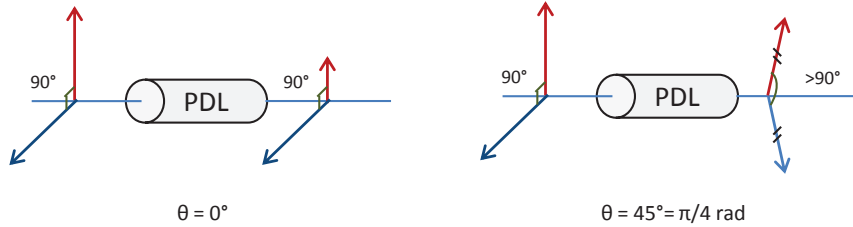


Figure 1.18: PDL effect: the reference signal without PDL propagates through a PDL element aligned with its polarization axes at left, and at an angle of $\pi/4$ with respect to its polarization axes at right.

Further findings on PDL

In [57], Huttner et al. prove that the interaction of PDL and PMD introduces anomalous dispersion and more severe effects than the ones produced if PMD or PDL are considered separately. The frequency dependence of PMD in (1.19) will turn the global PDL of the optical link into a frequency dependent variable. A given signal will observe a flat or a frequency-varying PDL depending on the signal bandwidth and the coherence bandwidth of the optical channel [58].

Moreover, field measurements of PDL in [52] showed that the probability distribution of Γ_{dB} was a truncated Gaussian distribution, rather than Maxwellian, because of the presence of a limited number of elements in the link having an appreciable PDL. Hence, there is no unique distribution to be considered when it comes to analyzing the penalties caused by PDL. However, in all cases, the global Γ_{dB} is a random variable that depends on the number of PDL elements in the link and also on their individual PDL values. As regards the temporal variations of PDL, they will follow the dynamics of PMD given that the instantaneous overall PDL results from the interaction of PMD with in-line PDL.

As mentioned earlier, PDL causes both a break of the orthogonality of the data-carrying polarizations as well as fluctuations of the OSNR on each polarization. Furthermore, at the dual-polarization coherent receiver, the OSNR per polarization depends on the accumulated ASE noise polarized parallel to the considered polarization state. The influence of ASE orthogonal to the signal is minimal as it does not coherently interfere with the optical signal. However, in the presence of severe in-line PDL, the distributed ASE noise emitted in the orthogonal polarization can become polarized parallel to the signal which also causes OSNR fluctuations [59, 60]. Interactions of PDL with non-linear effects were also studied in [61].

1.3.4.2 PDL management

While the dispersion compensation was thoroughly studied and completely compensated, PDL effects were not being considered until recently. System margins were left in order to absorb the PDL-induced penalties. Hence, PDL remains the main limiting linear impairment in PDM systems. In current systems (both with direct detection or coherent detection), PDL is not mitigated and only operating margins (up to 3 dB for PDM systems [59]) are considered to absorb the induced penalties and ensure a target performance. Obviously, in a power-efficient ultra-high bit rate communication scheme, system margins have to be chosen as small as possible and hence, PDL has to be managed via a specific processing.

The power fluctuations caused by PDL can be seen as a fading that cannot be undone by digital equalizers at the receiver. In some works, authors tried to reduce the penalties caused by PDL through whitening the ASE noise (de-correlating the noise on the x and y polarizations) [62] or by estimating the SOP of the signal and applying a certain channel inversion operation in Stokes space to cancel PDL [63]. However, in both cases, the penalty reduction was only partial. Moreover, the first method was effective only for specific rotation angles and the second method effective for low PDL values (up to 2 dB) and weak noise added at the transmitter side before the PDL element.

For finite PDL values, the 2×2 channel matrix is invertible and the transmission scheme has a MIMO structure as noted before. Hence, we will be interested in this thesis in applying more advanced MIMO digital signal processing to combat PDL. We will be presenting these MIMO

tools in chapter 2 and investigating their potential to mitigate PDL in PDM transmission systems in chapter 3.

1.3.5 Non-linear effects

The optical Kerr effect, in the propagation equation in (1.18), is driven by two components inducing a non-linear phase shift: the first is SPM as in (1.6) and the second is similar to XPM in WDM systems but results from the other co-propagating polarization at the same wavelength, and can be called intra-channel-XPM (IXPM). IXPM generates non-linear crosstalk between polarizations.

Moreover, WDM-PDM systems are impaired by another non-linear effect, known as cross-polarization modulation (XPolM) that also results from the optical Kerr effect. This time, the state of polarization (SOP) of a signal at a certain wavelength changes due to different non-linear phase shifts on its x - and y - components, induced by co-propagating channels at different wavelengths. This SOP variation depends on the intensity fluctuations of the WDM channels and can lead to depolarization, thus affecting polarization-sensitive receivers that may fail to track the non-linear SOP scattering. As for a WDM single-polarization system, a WDM-PDM system is engineered to operate on an optimal input power value that defines the NLT point for a specific set of parameters: the used modulation formats, the WDM channel spacing, PMD strength, CD dispersion map, the total length of the system as well as other parameters.

1.4 Space Division Multiplexing (SDM)

Space is currently considered as the last available degree of freedom that can be used in an optical fiber for increasing the capacity of the transmission system, since time, frequency, phase and polarization state were all investigated [64]. Being close to the capacity limit of an SMF-based system and with the growing and urgent need for higher capacities, it will become very difficult to achieve an increase in its spectral efficiency even if the most sophisticated optical amplification techniques, optical fibers, modulation formats and FEC codes are employed. Indeed, if we try to implement higher q-QAM modulations (such as 64-QAM) to achieve a high spectral efficiency over long distances, the span length has to be reduced and a large number of optical amplifiers used in order to limit the ASE noise power (as can be seen from (1.9)). Alternatively, the parallelism offered through the use of space as a multiplexing dimension can enormously reduce the cost and the energy consumption of the high-capacity system [65].

Spatial division multiplexing (SDM) is expected to offer high-capacity solutions that would be more power-efficient and have a lower cost than the simple deployment of more single mode fibers in core networks (metropolitan, regional and long-haul). SDM is also an interesting technology for-short distance ultra high-capacity networks such as data center interconnections. “In-fiber”

SDM can be implemented using multi-mode fibers (MMF) or multi-core fibers (MCF) to better use the space inside the fiber [1]. In the former case, the core of a fiber is enlarged to allow the propagation of more than one spatial mode [9]. Each mode describes a spatial distribution of the transverse electrical field and the modes form an orthogonal set that can be used for multiplexing. In the latter case, multiple cores are inserted in the cladding with each core being single-mode or multi-mode [66]. Different SDM technologies are shown in Fig. 1.19.

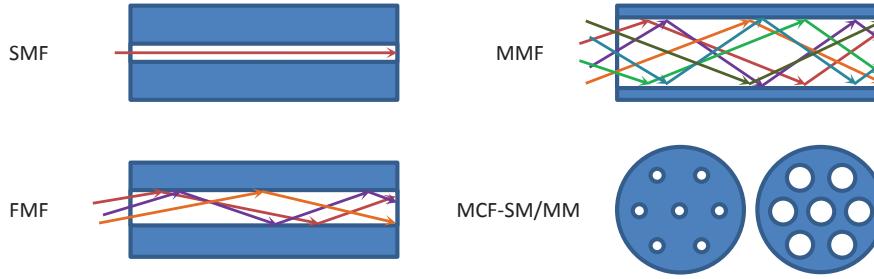


Figure 1.19: The different spatial multiplexing technologies: multi-mode, few-mode fibers, and multi-core fibers where cores can be single-mode or multi-mode.

The key advantage of these SDM techniques is the provision of an orthogonal set of pathways that can carry independent data streams in the same fiber while answering integration constraints of system components, thus reducing the cost and the energy consumption of the whole system, such as the integration in the fiber itself [10, 13], at the transponders [67] and also at the optical amplifiers [68]. By increasing the density of spatial channels (modes and/or cores) in the fiber, SDM solutions prevent the need to install parallel fibers and multiple components which suggest a linear increase of the power consumption and system cost with an increasing capacity. However, integration will introduce crosstalk between the channels which can severely limit or prohibit the expansion of SDM systems unless that crosstalk is addressed in-line or digitally at the transceivers.

1.4.1 Multi-core fibers

The concept of SDM in the fiber goes back to 1979 when multi-core fibers were suggested to multiplex data [69]. Yet, no further investigations were carried on multi-core or multi-mode fibers for long-distance communication, mainly due to their high losses and crosstalk levels. These fibers were rather used as sensing devices and for laser fabrication.

In MCFs, the crosstalk increases with a smaller inter-core distance. At the same time, for fiber reliability issues, fiber cladding diameters of over 200 μm are not allowed which places a limit on the number of integrated cores. MCFs with 7 to 19 cores [70] placed on an hexagonal grid were designed to allow for an optimal and compact placement where all the nearest-neighbor

cores are separated by the same distance as can be seen in Fig. 1.19. In [71], a 52-km transmission experiment using a 7-core fiber was demonstrated with a single-channel rate of 456 Gb/s. It consisted of 222 WDM channels in each core resulting in a net total capacity of 1 Pb/s.

Depending on the fiber design and specifically the inter-core distances, we can speak of independent-cores MCF with uncoupled parallel channels [72] or coupled-core MCFs where the data launched in one-core leaks into the others [73]. In the second case, a better spatial density is achieved but the transmitted data has to be recovered through MIMO techniques at the receiver after the coherent detection.

1.4.2 Multi-mode fibers

When a fiber contains a single large core, we speak of multi-mode fibers (MMFs), and if the spatial modes are between 2 and 10, we refer to them as few-mode fibers (FMFs). The number of guided spatial modes increases rapidly with the core diameter and is determined through the normalized frequency V , also called V -parameter, that depends on optical and geometrical characteristics of the fiber [19, Chap.2]. For step-index fiber, V is given by (1.3). Like all electromagnetic phenomena in a waveguide, the propagation of an electrical field in a fiber is governed by Maxwell's equations. Using a weakly-guiding assumption for MMFs, the solutions to these equations give rise to exact modes that can be combined into M linearly polarized (LP) and mutually orthogonal spatial modes [18, Chap.4], each supporting two polarization modes.

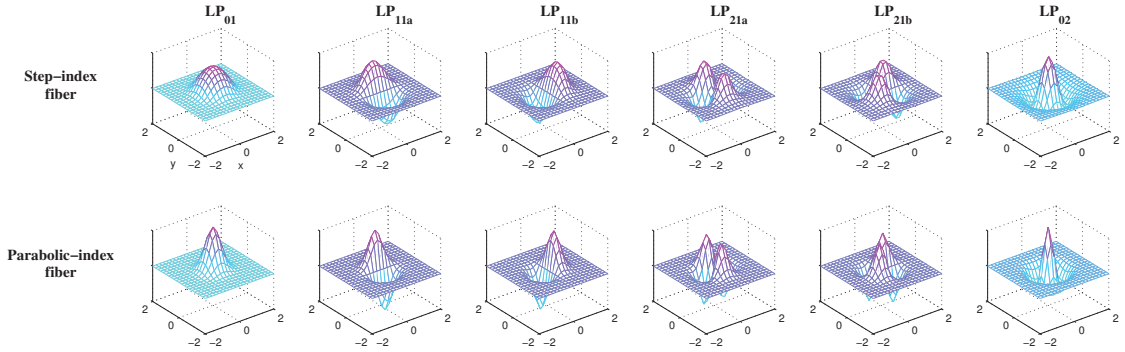


Figure 1.20: Spatial distributions of the first six modes given by the linearly polarized (LP) mode representation for a step-index fiber (first row) and the Laguerre-Gauss approximation for a graded-index fiber (second row). The (x, y) grid goes from $-2r_c$ to $2r_c$.

The spatial distributions of the fields of LP modes are approximated by Bessel functions in fibers having a step-index profile. For fibers with a parabolic-index profile, a sub-category of gradient-index fibers, a Laguerre-Gauss (LG) representation of the LP modes gives a better approximation. However, both representations show similar intensity profiles as can be seen in Fig. 1.20 for two 6-mode fibers with the same core radius and different index profiles. Using the LP-mode representation, a mode is denoted $LP_{\mu\nu}$ where μ is the azimuthal order and ν is the

radial order of the mode. For higher-order azimuthal modes ($\mu \geq 1$), the same spatial distribution can have two possible orientations, resulting in mode degeneracy. Suffixes a and b are then used to differentiate the modes. For a given index profile of the fiber, we can draw cut-off diagrams as a function of the V -parameter in order to determine the number of propagating modes for a given value of V [18, Chap.4].

Mode division multiplexing (MDM) is realized by sending independent data streams on each of the M modes with the intent to increase the capacity by M compared to single-mode fiber based schemes. In MMFs, the spatial modes overlap significantly and the carried data streams couple randomly during the propagation. Moreover, the modes have different propagation velocities and will exhibit differential modal group dispersion (DMGD). These disparities in velocities are smaller in fibers with a parabolic-index profile compared to step-index fibers. The modes can also experience differential losses or gains leading to loss disparities and a reduction of the expected capacity [64, 74]. In the next sections, we will focus on these effects especially in FMF-based coherent SDM systems that are rapidly gaining maturity. Once the design of MMFs and single-mode or multi-mode MCFs becomes more mature and practical, we will have to deal with similar crosstalk issues arising from the fiber and the optical components.

It is also important to remind the reader that only coherent-detection MDM systems will be considered because direct-detection systems would require stringent crosstalk and distortion conditions such that the modes can be efficiently separated at the receiver [75]. On the other hand, with coherent detection, the transmitter needs to multiplex the signals into an orthogonal set of modes and the receiver needs to project them on an orthogonal set, then MIMO DSP will separate the mixed signals in the electrical domain. The architecture of an MDM coherent receiver is a generalization of the dual-polarization coherent receiver in Fig. 1.14 with the addition of mode de-multiplexers to separate the modes.

1.4.3 Linear modal crosstalk

The linear coupling or crosstalk between the different spatial modes strongly depends on the fiber design parameters, the fabrication methods and environmental stresses. Hence, various amounts of linear crosstalk between the different pairs of modes have to be considered when studying the performance of an MDM system. The propagating modes spatially overlap in the fiber due to imperfections in doping and geometrical asymmetries caused by mechanical or thermal stress and micro-bends. The coupling strength also depends on the values of the propagation constants β of the modes: closer values induce more crosstalk. This overlap consists in an energy transfer between modes. The symbols carried by each mode randomly and ceaselessly couple along the fiber resulting in distributed crosstalk, but also at discrete points causing discrete crosstalk (splices, connections, in-line components such as multiplexers and de-multiplexers). This coupling can be

unitary and lossless as considered in [76, 77] or non-unitary and lossy, and introduces in this case loss disparities, as in [78] depending on the nature and the sources of imperfections. In all cases, it can be represented by an $2M \times 2M$ coupling matrix $\mathbf{C}$ (M spatial modes with 2 polarizations per mode). When the crosstalk is unitary, the corresponding matrix obeys the following property: $\mathbf{C}\mathbf{C}^\dagger = \mathbf{I}$. We will look into different coupling scenarios while studying MDM systems in chapter 4.

Two coupling strategies were recently investigated: some works have considered reducing the modal coupling in FMFs [79, 80], keeping the modal crosstalk levels very low to avoid the use of a full MIMO processing at the receiver. These works considered coupling as a limiting factor of MDM transmissions just as ASE noise and found the maximum allowed crosstalk levels for the different modulation formats [81]. On the other hand, other works considered that coupling is inevitable and would be generated at the optical components even if it is kept low in the fiber itself, hence a full MIMO processing is needed [82, 83] and the MIMO algorithms used for $2M \times 2M$ channels being simply a generalization of the 2×2 algorithms in PDM systems.

Furthermore, Kahn et al. [74, 84] suggested to positively exploit strong coupling in order to decrease the differential modal losses as well as the maximum modal delay spread induced by DMGD, that will follow a random-walk process in the strong coupling regime and thus scale with the square-root of the fiber length rather than scaling linearly in the case of weak coupling (just like PMD and random polarization rotations in PDM systems).

1.4.4 Differential Modal Group Dispersion (DMGD)

Each spatial mode propagates with a different velocity, resulting in differential mode group delays $\Delta\beta$ in the fiber. This difference is intrinsic to the fiber and exists even in a perfectly symmetric fiber. It can be evaluated knowing the index profile of the fiber. For a step-index profile, the delays can become extremely larger for higher order modes, thus inducing ISI of several consecutive data symbols and prohibiting the simultaneous processing of all the modes at the receiver [85]. For a parabolic-index profile, the modes can be separated in modal groups with a constant inter-group delay [18, Chap.5].

As for PMD in SSMFs, DMGD in MMFs can be modeled by a concatenation of uncoupled pure modal dispersion elements $\mathbf{D}_k$ and coupling elements $\mathbf{C}_k$. $\mathbf{D}_k$ is an $2M \times 2M$ diagonal matrix with different phase elements:

$$\mathbf{D}_k = \begin{pmatrix} \exp(-i(\phi_{1,k} + \omega\tau_{1,k})) & & 0 \\ & \ddots & \\ 0 & & \exp(-i(\phi_{M,k} + \omega\tau_{M,k})) \end{pmatrix} \quad (1.24)$$

where $\phi_{i,k}$ is a random phase shift accounting for modal noise (random phase differences between the spatial modes) and $\tau_{i,k}$ is the uncoupled group delay in ps of the i^{th} mode at the k^{th} fiber element.

Realizing fibers with very low DMGD for several spatial modes is a very challenging task. Recently, important advancements in the fabrication of few-mode optical components such as few-mode fibers with low DMGD [9, 10] obtained by adding a trench in the index profile of the cladding, in order to reduce the difference in the group velocities as well as the coupling between the propagating modes, led to the experimental demonstrations of 3-mode and 6-mode MDM systems.

1.4.4.1 DMGD management

A substantive reason for the current interest in MDM systems is the integration of coherent detection and DSP techniques to compensate various impairments in the optical networks. As a result, unitary modal crosstalk can be managed electronically at the receiver after the coherent detector transposes the optical signal into the electrical domain.

Differential propagation delays along with crosstalk lead to temporal and spatial ISI. At a given time slot t , the data symbols carried by M propagating modes will spatially mix with each others as well as with the symbols sent at the previous T_{mem} time slots corresponding to the maximum delay caused by DMGD, also called the memory of the channel. To completely compensate for unitary crosstalk and ISI, the equalizer at the receiver should be able to address this maximum delay spread between the propagating modes [83]. The computational complexity of several time-domain and frequency-domain equalizers in both single-carrier and multi-carrier formats was studied and OFDM was found to achieve the lowest complexity for long-haul MDM systems [86]. When an OFDM modulation is considered, a single-tap frequency-domain $2M \times 2M$ filter, per subcarrier, is enough to decouple the modes provided that a cyclic prefix longer than the maximum dispersion spread is inserted to absorb ISI. More details on both equalizing techniques will be given in the next chapter.

For all equalizers, the complexity and/or the overhead are preferred to be small. Hence, the maximum delay spread in thousand-kilometers-long transmissions, should be kept as low as possible. For the moment, this is possible for few-mode fibers (FMFs) only. Graded-index core fiber designs [10] can achieve a DMGD as low as 40 ps/km [87] for a 6-mode fiber supporting the following linearly polarized modes: the fundamental LP_{01} mode and the LP_{11a} , LP_{11b} , LP_{21a} , LP_{21b} and LP_{02} modes. Another possible way of limiting the delay spread is to combine fibers with DMGD of opposite signs, offering DMGD cancellation and resulting in dispersion values as low as 6 ps/km [88] for a 3-mode fiber (the fundamental LP_{01} mode and the LP_{11a} , LP_{11b} modes). Distances of more than 1000 km were achieved with 3-mode FMFs [82] and successful transmission

experiments over 130 km were conducted with FMFs supporting 6 modes [89]. However, scaling beyond 6 modes while maintaining low DMGD values is currently very challenging.

In all cases, strong coupling should be seen as an advantage because it reduces the maximum delay spread and thus the signal processing complexity [74, 84]. If the crosstalk in the channel was unitary, meaning that the paths remain orthogonal after propagating through the optical channel, separating them and removing ISI induced by the differential delays, would be enough in order to retrieve the performance of a crosstalk-free channel. However, long-haul MDM systems are impaired by modal loss disparities that induces a non-unitary effect known as mode dependent loss (MDL).

1.4.5 Mode Dependent Loss (MDL)

In long-haul MDM systems, MDL arises from imperfections in the link such as non-unitary couplings in the fiber due to bending losses and splices [2, 90, 91] as well as mode dependent gains at optical amplifiers [68, 92] and other MDM components (couplers, multiplexers...). Indeed, the architecture of multi-mode amplifiers must be adapted to the propagating spatial modes in order to provide similar gains for all modes. To do so, the transverse intensity pattern of the pump and/or the Erbium distribution in the core of the amplifying fiber must be tailored for the specific modes. Another investigated solution consists in micro-structuring the fiber core of the EDFA. However, in all cases, zeroing the mode dependent gain is not possible [92]. Unlike dispersion, MDL can reduce the capacity of the MIMO channel [64, 74] because the propagating modes will lose their orthogonality and will observe various attenuation levels. Some may have very low power levels at the receiver, which translates into a low received OSNR [2, 90]. Thus, it is crucial to maintain MDL levels as low as possible in the link.

In [78], Juarez et al. showed that the fiber-generated MDL due to micro-bends is limited to 2 dB if the overall coupling loss is lower than 1 dB in a single kilometer of fiber supporting a large number of modes. On the other side, it is quite difficult to achieve similar low levels of gain offsets in the optical components such as in FMAs due to technological limits (doping profiles in optical amplifiers, spatial distribution of the pumping power...). In [91], the authors measure a gain difference of 18 dB between the most privileged and the least privileged modes in a 6-mode SDM system made of 10 spans of 50km FMFs. Recent advancements in the design of Erbium-doped amplifiers [92] succeeded to provide FMAs with a minimum modal gain offset of 1.3 dB between the most and the least amplified modes for a 3-mode FMA, and 2 dB between the most and the least amplified modes for a 6-mode FMA. Knowing that optical amplifiers are periodically inserted in a metro or long-haul optical link, the gain offsets will lead to important overall MDL levels. Moreover, in [93], through an experimental study of the transfer matrix of a 55-mode fiber,

the authors point out that MDL steadily increases for higher order modes and the orthogonality of modes cannot be maintained due to imperfect optical components.

1.4.5.1 MDL management

Simulation studies were carried out to reduce MDL in long-haul optical links where FMAs are periodically inserted, using fibers with strongly coupled modes [2] and mode scramblers that intentionally couple the modes [90, 94] in order to average the losses experienced by each mode and thus reduce the accumulated MDL. In [74, 84], the authors have also shown the benefits of strong mode coupling in reducing both DMGD and MDL.

Moreover, in [2], Lobato et al. have tested the performance of different equalization techniques for an MDL-impaired OFDM-MDM system with channel state information available at the receiver. They showed that the use of an optimal, yet more complex, maximum-likelihood (ML) detection instead of a zero-forcing (ZF) equalizer enhances the performance of the system, which is expected in presence of MDL due to the channel inversion operation of the ZF equalizer that enhances noise and leads to a sub-optimal detection. In [94], the same authors proposed a reduced-complexity sub-optimal version of the ML detection, called the reduced-search ML decoder. More details on the different equalization techniques as well as their advantages and drawbacks will be exposed in chapter 2. In Fig. 1.21 taken from [2], we show the achieved reductions of penalties induced by MDL realized through the use of strong-coupling fibers and ML decoding.

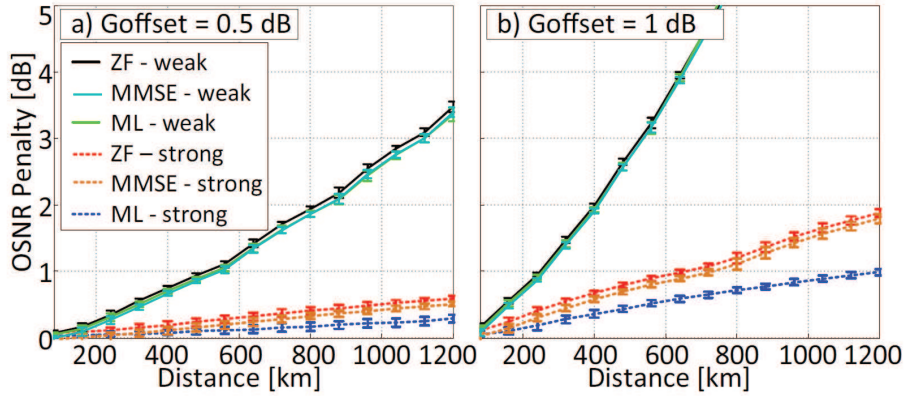


Figure 1.21: Average and standard deviation of the OSNR penalties at $\text{BER} = 10^{-3}$ for ZF-, MMSE- and ML- decoded 3-mode transmission impaired by MDG at amplifiers (for two different gain offsets between LP_{01} and the LP_{11} modes) placed each 80 km of fibers in the weak and strong coupling regimes. Figure taken from [2].

However, the proposed coupling and scrambling solutions have their limitations and can set some stringent requirements in order to obtain the desired MDL reduction (fiber design, number of required scramblers). Indeed, the proposed mode scrambling solutions remain conceptual and

mode scramblers are simulated as loss-free perfect mode permutations [90] or inherently strong-coupling fiber sections [94] while real mode scramblers would suffer from insertion losses and lossy crosstalk. On the other hand, in [2], ML detection showed a slight performance enhancement compared to ZF equalization only when modes are strongly coupled. Thus, MDL mitigation remains an open problem and more sophisticated MIMO tools are needed to mitigate MDL.

1.4.6 Non-linear effects

As can be seen in the previous sections, numerous linear effects have to be addressed in order to guarantee that SDM is a viable plan for future optical transmission systems. However, the design of an optical link must take into account the non-linear effects that arise in the fiber when the injected input powers are sufficiently high or the transmission distance is sufficiently long. In our work, we will solely focus on linear effects in MDM systems.

Knowing that the effective core areas of multi-mode fibers are larger than those of SMFs, the non-linear parameter γ_{NL} given in (1.13) which is inversely proportional to A_{eff} will be smaller for MMFs and one would expect an enhanced tolerance to non-linearity. Moreover, in a noteworthy work [95], the authors theoretically studied inter-modal cross phase modulation and inter-modal four-wave-mixing in MDM systems. They found that rapidly varying birefringence in the fiber reduces the number of inter-modal non-linear terms in the NLS propagation equation which can be written for the p -th mode as [96]:

$$\frac{\partial \mathbf{a}_p}{\partial z} = -\frac{\alpha}{2} \mathbf{a}_p - \beta_{1p} \frac{\partial \mathbf{a}_p}{\partial t} - \frac{i\beta_{2p}}{2} \frac{\partial^2 \mathbf{a}_p}{\partial t^2} + i \frac{\gamma_{NL}}{3} \sum_{l,m,n} f_{plmn} [(\mathbf{a}_l^T \mathbf{a}_m) \mathbf{a}_n^* + 2(\mathbf{a}_n^\dagger \mathbf{a}_m) \mathbf{a}_l] + i \sum_m q_{mp} \mathbf{a}_m \quad (1.25)$$

where $\mathbf{a}_p = [a_p X, a_p Y]^T$ is the electrical field vector of the p -th mode, f_{plmn} and q_{mp} are the non-linear and linear couplings among spatial modes. However, further investigations of non-linear effects in multi-mode and multi-core fibers are required such as the influence of discrete non-unitary couplings as noted in [97] in order to establish the capacity limit for SDM schemes as well as their NLT operation point.

1.5 Beyond 400G: WDM-SDM-PDM super-channels

The standard 100 Gb/s channel modulation in the currently available commercial systems is based on single-carrier PDM-QPSK. Field trials of WDM-PDM systems with a dual-carrier channel modulated with 16-QAM symbols have been recently demonstrated [98] yielding a net bitrate of 400 Gb/s per wavelength over a 550-km-long commercial system. However, long-distance Terabit single-wavelength rates seem extremely difficult to achieve using only high order modulations and polarization multiplexing. For instance, doubling the spectral efficiency of 8 bits/s/Hz in [98]

would require a 256-QAM constellation over each polarization. The increased noise sensitivity of this constellation would severely reduce the capacity-distance product of the transmission link.

It is unclear whether 400 Gb/s or 1 Tb/s will be the next channel capacity standard beyond 100 Gb/s but either way, new channel designs are needed to achieve reasonable transmission distances. On one hand, due to the bandwidth limitation in electronic components, the symbol rate cannot increase much and on the other hand, if we increase the modulation level, the transmission distance will have to be reduced. Therefore, 100-200 Gb/s systems may be the last high-capacity single-optical-carrier channels and future systems would require more than one optical carrier. For instance, a 1 Tb/s channel could be a WDM “super-channel” made of five PDM-16-QAM 200 Gb/s carriers.

The sub-channels of a super-channel can be packed in several different ways and tighter than the conventional 50 GHz spaced channels because they all belong to the same channel. They can consist of single-carrier orthogonal channels, called all-optical OFDM [99] with an adjacent carrier spacing equal to the symbol rate to grant orthogonality. Another flavor of super-channels is known as Nyquist WDM where the carriers have a rectangular spectrum and a bandwidth equal to the symbol rate [100]. A super-channel can also be obtained by forming a multi-band OFDM signal where each sub-channel consists of a conventional OFDM signal [101]. Every super-channel technology has a number of pros and cons, however they all provide an enhanced occupation of the spectrum compared to conventional systems provided that a flexible spectral grid can be adapted with a granularity smaller than 50 GHz. However, the drawback of introducing flexible spectral bandwidth is the increase of network operational complexity (the need for more accurate optical switches).

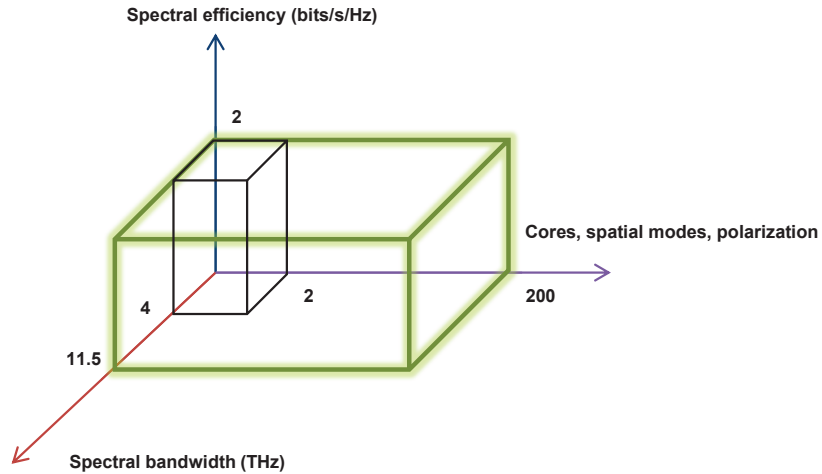


Figure 1.22: Key parameters for the capacity increase of optical fiber transmission system: current single-wavelength 100G systems in black and proposed hypothetical system in green.

Opening new optical amplification spectral windows in the S-band (for short- λ band) or in the L-band is also considered as a candidate solution to increase the capacity of future optical transmission systems. However, many technological challenges would limit the capacity gain such as higher noise figures in the new windows and gain fluctuations among the channels. Hence, any important capacity leap will have to use space as a new degree of freedom.

Space division multiplexing is emerging as an interesting candidate for the increase of the capacity as well as the capacity-distance product of optical communication systems. A number of challenges are being addressed such as the large modal delays, modal crosstalk, MDL and the scaling of SDM components. The diversity of SDM techniques (MMFs, MCFs or a combination of both) makes this multiplexing solution a very interesting research topic. A technological breakthrough offering a minimum of 10-fold increase in capacity is expected of SDM. The proposed solution should also have a cost and energy consumption advantage over the deployment of a single-mode-fiber based infrastructure, thus reducing the cost-per-bit of the system.

SDM is closer to PDM than to WDM because wavelengths are able to share the optical components (amplifiers, fibers,...) without severe crosstalk while SDM requires special designs of integrated components where crosstalk has to be managed. Given the urgent need for higher capacities, major research works are investigating the feasibility of SDM, trying to surmount its engineering problems. However, it should not be denied that spatial multiplexing provides an opportunity to design high-capacity, flexible and scalable WDM-SDM-PDM super-channel systems. One can imagine a system made of 10-core fibers where each core is multi-mode with 10 propagating polarization-multiplexed spatial modes, and fill 225 channels in the C- and L- bands with a spectral granularity of 50 GHz. Hence, sending a 4-QAM modulation on each of the 200 spatial and polarization modes with a per-channel symbol rate of 28 GBaud, we can achieve a bitrate of 10 Tb/s per wavelength and a total bitrate of 2.25 Pb/s per fiber. This hypothetical system is represented in Fig. 1.22 and compared to a typical currently available 100G system of 80 to 100 wavelengths. Depending on whether the suggested scheme is proposed for long-distance transmission systems or short-distance interconnections with less ASE and non-linear noise accumulation, the modulation format can be changed to enhance the spectral efficiency.

Summary

In this chapter, we reviewed the evolution of optical fiber transmission systems leading to the emergence of high-capacity MIMO PDM and SDM schemes that increase both the reach and the capacity of the link. The major impairments affecting these systems are:

1. Linear differential dispersion effects : polarization mode dispersion (PMD) and differential modal group dispersion (DMGD) inducing inter-symbol-interference.

2. Linear crosstalk : polarizations and modes spatially overlapping all along the fiber and in the optical components.
3. Linear differential losses : polarization- and mode- dependence of optical components due to challenging technological limits. Hence, the channels experience unequal losses when propagating through the various components distributed in the optical link. These losses accumulate leading to important disparities between the energies of the different channels and a loss of their orthogonality.
4. Non-linear crosstalk : cross-phase/-polarization/-mode modulation (XPM) are non-linear effects impacting the multiplexed systems and limiting the power levels that can be injected in the optical fiber.

In the next chapter, the focus will be placed on the digital signal processing algorithms that were implemented in the deployed systems as well as on new MIMO techniques that can further enhance the performance of current and future optical transmission systems.

Chapter 2

Advanced Signal Processing and Coding Techniques for Optical Communications

We have listed in the previous chapter the main physical effects that impair optical fiber transmission systems. We have also mentioned whether these effects were optically or digitally managed or not. Coherent detection, fast analog-to-digital converters [102] and high-speed integrated circuits such as field-programmable gate arrays (FPGA) or application-specific integrated circuits (ASIC) pioneered the introduction of more and more sophisticated digital signal processing algorithms initially designed for wireless communication systems, in optical transceivers. The ultimate goal of DSP and coding is to bring the system performance closer to its Shannon capacity limit, or more specifically to the non-linear Shannon capacity limit of optical fiber communications [6].

The first DSP tools were meant to increase the robustness of the optical system against additive ASE noise and dispersion effects. Forward error correction (FEC) codes and compensation of deterministic chromatic dispersion were the first techniques to be introduced. Later on, with the appearance of MIMO schemes through polarization division multiplexing, adaptive MIMO equalizers made their way into the deployed coherent receivers to mitigate random effects (PMD, non-linear effects...). Along with these main algorithms, frequency and time synchronization, carrier phase estimation and non-linear compensation techniques were also employed to enhance the reach and the quality of the transmission system.

The most important advantage of coherent detection is probably the access to all the characteristics of the optical signal and hence the ability to modulate and detect advanced modulation formats from single-carrier higher-order digital modulation formats (PSK, QAM) that are currently used in commercial systems to multi-carrier analog formats such as orthogonal frequency division multiplexing (OFDM) that has proven to be as effective as single-carrier based schemes

in mitigating dispersion effects in optical MIMO channels [103] with a complexity advantage particularly notable for SDM systems [104].

Furthermore, emerging MIMO schemes pave the way to a new class of modulation symbols that does not use the MIMO structure only for multiplexing, but also exploits it to protect the transmitted information from non-unitary channel effects. These new symbols are called “Space-Time (ST) codes” because they were designed for multi-antenna wireless systems and employ the spatial and temporal dimensions to protect the data. In 2010, our research group was the first to apply these codes over the two dimensions: polarization and time, for PDL mitigation in PDM systems [8], a capacity-limiting channel impairment that was left uncompensated for. Building upon these observations, we are interested in investigating the capabilities of ST coding in larger optical MIMO systems. We shall see that even though MIMO processing already made its way into optical communications, the MIMO equalizers implemented in today’s optical transceivers achieve only a basic processing and the potential of MIMO DSP is far from being exhausted whether we consider the transmitter side or the receiver side.

In this chapter, we focus on the DSP algorithms and coding schemes that have been applied in optical communication systems to process the received data and correctly detect the binary information. We start by the channel codes or FEC codes that enable the correction of bit errors by encoding the information bits at the transmitter. Then, we describe the equalizers used for dispersion compensation and separation of the mixed channels (polarizations in PDM systems or both modes and polarizations in SDM systems). We focus in particular on frequency-domain equalization along with the multi-carrier OFDM modulation format that opens up new coding and modulation opportunities for emerging optical MIMO schemes. Finally, we present the principle of ST coding as well as the major codes that were designed for the wireless channels and will be applied to the optical channels in the following chapters, then end up with a review of MIMO decoders.

2.1 Forward Error Correcting (FEC) codes

Long-haul optical transmission systems require a high-quality bit error rate (BER) performance. In optical communications, a transmission can be considered as error-free if the BER is lower than $10^{-13} - 10^{-15}$ which implies a very high OSNR at the receiver and a perfect compensation of the various channel effects. These conditions are not easily met over long distances. Therefore, forward error correction (FEC) codes were introduced to enable the correction of bit errors - mainly induced by ASE noise, modeled as an additive white Gaussian noise - by encoding the information bits at the transmitter. The coding gains brought by FEC codes offer an improved OSNR sensitivity to the installed optical systems, hence allowing to upgrade their capacity (by adding more WDM channels with a lower power per channel) and to increase their transmission reach.

2.1.1 Basics of FEC

A FEC code introduces redundancy to the transmitted data, referred to as FEC overhead, in order to protect the information and reduce the BER lowering it from a high pre-FEC BER value to a low post-FEC value. The encoding rule is known at both the transmitter and the receiver. The redundancy increases the total bit rate, however the provided protection results in a coding gain evaluated as the difference in the required OSNR between the FEC-coded and the uncoded schemes. The maximum attainable net coding gains (obtained from the coding gains and taking into account the added redundancy) are given by the Shannon limit [105].

Linear block codes are a FEC family where information is encoded in blocks. A given code is characterized by its rate r_{FEC} defined as the ratio between the number of information bits k and the number of transmitted bits n (containing the useful data + redundant parity bits):

$$r_{FEC} = \frac{k}{n} \quad (2.1)$$

Let E_b denote the information bit energy, E_c the transmitted bit energy and E_s the modulated symbol energy chosen from a q -point constellation. Since the total information word energy kE_b must be the same as the total codeword energy nE_c , we obtain the following relation between E_b and E_s :

$$E_s = \log_2(q)E_c = \log_2(q)r_{FEC}E_b = k_cr_{FEC}E_b \quad (2.2)$$

The encoding operation corresponds to the product of a k -bit data vector with a FEC generator matrix that defines the encoding rule. The n -bit codeword is then modulated and sent through the channel. At the receiver, after demodulation, the n -bit sequence might no longer be a codeword due to the propagation through a noisy channel. The decoder tries to estimate the original data by comparing the n -bit received vector with the possible transmitted codewords. It decides in favor of the codeword that is closest to the received vector (by closest, we mean the one that has the lowest number of different bit positions with the received vector). This decoding rule is known as maximum-likelihood (ML) decoding and results in a minimum probability of a codeword error for the binary symmetric channel [11, Chap.7]. It can be implemented in several ways depending on the properties of each code family.

FEC decoders can be generally categorized as hard- or soft-decision decoders, depending on whether they apply a per-bit threshold decision where each sample corresponding to a single bit is quantized to 0 or to 1, or whether no quantization (in practice, a less severe quantization than the binary one) is performed and a probabilistic approach that exploits the available redundancy is used to correct the errors and estimate the information bits. In optical systems, the soft-decision quantization can be performed with an 8-bit resolution provided by powerful ADCs optimized for optical transmission systems and operating at 63 GS/s or even 90 GS/s [56, 102].

An important characteristic of a code is its minimal distance $d_{min,FEC}$ defined as the minimal number of distinct bits between any two different codewords. This minimal distance quantifies the error detection and error correction capabilities of a code. Indeed, a linear block code denoted by $\mathcal{C}(n,k,d_{min,FEC})$ can correctly detect errors when the number of errors t_d is strictly less than $d_{min,FEC}$. In order to correct errors, the decoder has to estimate the most likely codeword. If there are more than $d_{min,FEC}/2$ errors, the estimated codeword will not be the emitted one, thus the maximum number of errors t_c that a code is able to correct is: $t_c = \lfloor \frac{d_{min,FEC}-1}{2} \rfloor$.

2.1.2 FEC in optical transmission systems

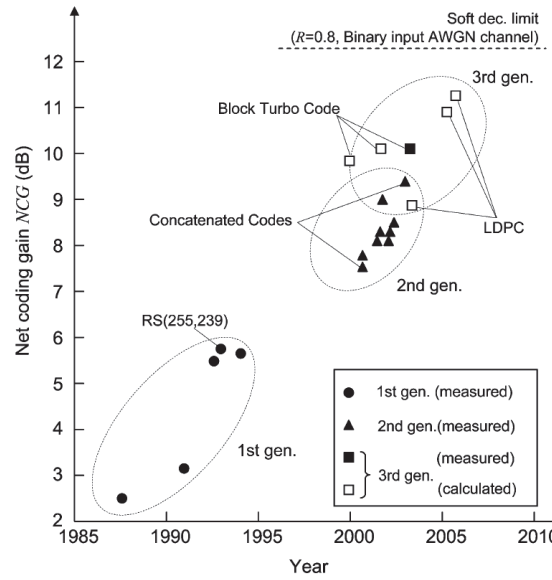


Figure 2.1: Recent progress in FEC for optical communication systems taken from [3]. AWGN stands for Additive White Gaussian Noise.

FEC coding is needed to enhance the receiver sensitivity especially when moving to higher order modulation formats (Appendix A). The most suitable FEC codes for optical communications should have the following properties: limited overhead ($1 - r_{FEC}$) (typically 7% but can reach 25% in submarine ultra long-haul systems in which important FEC coding gains are needed in order to ensure a highly reliable transmission), short to medium codeword length ($n < 10000$ bits), moderate coding/decoding complexity and more stringent transmission quality requirement than wireless channels (low post-FEC BER, typically 10^{-13}). The highest pre-FEC BER value that can be corrected to a BER lower than 10^{-13} is referred to as the FEC limit that we show in Fig. 1.10.

FEC has been implemented in optical transmission systems since the early 1990s. First-generation FEC used linear block codes such as Reed-Solomon (RS) codes along with hard-decision decoding (HDD), in particular RS(255,239) with a code rate $r_{FEC} = 0.93$ was specified in the ITU-T recommendation G.975 and provided a net coding gain of 5.8 dB at BER =

10^{-13} . Later, as WDM matured, more powerful FEC codes were needed to improve more the OSNR sensitivities. Hence, second-generation FEC consisted in two concatenated codes chosen from the Reed-Solomon and Boss-Chaudhuri-Hocquenghem (BCH) code families with code rates $r_{FEC} = 0.79 - 0.93$ along with an intermediate interleaving block. Interleaving data enhances the robustness of the transmission system to error-bursts that can be induced for instance by PMD in direct detection systems or non-linear effects such as XPM [106]. Using the RS-BCH interleaved scheme, the achieved net coding gains increased to 7 – 9 dB.

More recently, the interest shifted to more advanced FEC schemes with higher net coding gains to achieve both a higher capacity and a reduction of the upgrade costs of the optical transmission systems (minimizing the number of optical amplifiers, using legacy fibers...). To answer these needs, linear dispersion parity check (LDPC) codes invented by Gallager [107] along with soft-decision decoding (SDD) were investigated [108, 109] on 100 Gb/s coherent single-carrier and OFDM systems, bringing more coding gains than the previous generations. This soft-decoded LDPC scheme belongs to a new class, the third-generation FEC in optical systems consisting of modern coding techniques such as turbo codes and product codes along with iterative SDD, code rates $r_{FEC} > 0.8$ and net coding gains exceeding 10 dB. However, LDPC codes showed lower decoding and implementation complexity and better performance compared to turbo codes and product codes [3, 110]. One drawback of LDPC codes is the presence of error floors in their BER performance that appears around $BER = 10^{-10}$. Among the proposed solutions to remove the error floor with expected coding gains exceeding 10 dB, a concatenation of an LDPC code with an RS code was demonstrated in [111] and new LDPC constructions were suggested in [112].

For the recently developed 100 Gb/s PDM-QPSK coherent transceivers, FEC was integrated with the remaining DSP blocks on a single ASIC to enable the implementation of soft-decision decoding. SDD-FEC is a key requirement for the realization of reliable WDM long-haul and ultra-long-haul transmissions with more than 100 Gb/s rate per wavelength [56]. The powerful SDD-FEC scheme has a number of useful applications in future optical transmission systems such as the relaxation of the quality requirements of various optical components and the ability to lower a pre-FEC BER of $10^{-2} - 10^{-3}$ to a post-FEC BER of 10^{-15} or even lower with a 15% overhead [3, 56]. The evolution of FEC coding solutions in optical communication systems is shown in Fig. 2.1.

Beyond FEC coding, new coding approaches combining channel coding with the modulation process are also studied in order to improve the performance, the spectral efficiency and the reach of the system while maintaining low symbol rates and net data rates of 100 Gb/s per wavelength. These approaches include multi-dimensional constellations associated with set partitioning or bit-interleaved coded modulations [113]. As in FEC coding, a redundancy is added in order to better protect the transmitted information. The obtained performance gain stems from the larger minimal

distance between the constellation points compared to the common 2-dimensional QAM constellations, leading to improved OSNR sensitivity limits and more power-efficient formats for optical communication systems [39].

2.2 Basic digital equalization in 100 Gb/s single-carrier PDM-QPSK systems

The diagram in Fig. 1.12 shows the main functional DSP blocks at the receiver side for coherent optical communication systems. Their aim is to ideally compensate or partially mitigate the various linear and non linear, deterministic and random, impairments of the channel. In this section, we focus on the DSP tools that followed the deployment of dual-polarization coherent receivers in optical communication systems. The corresponding transmitter and receiver are given in Fig 1.14. The photodetectors at the receiver side give access to the complete information on the propagating optical field in the electrical domain (amplitude, phase, polarization).

After analog-to-digital conversion, sampling at a frequency $f_e = 1/T_e$ and normalization, the received signals at time k are given by:

$$\begin{aligned} y_X(k) &= I_X(k) + iQ_X(k) \\ y_Y(k) &= I_Y(k) + iQ_Y(k) \end{aligned} \quad (2.3)$$

For a complete received frame, the complex digitized vectors $\mathbf{y}_X = [y_X(k), k = 1 \dots K]$ and $\mathbf{y}_Y = [y_Y(k), k = 1 \dots K]$ are processed to compensate for chromatic dispersion, polarization mode dispersion and crosstalk. Then, the frequency offset between the signal and the local oscillator is removed as well as the phase noise resulting from lasers at the transmitter and the receiver before demodulating the signal and decoding the binary information. Apart from these steps, fiber non-linearity compensating equalizers may be performed in conjunction with CD compensation, however we will solely focus in this thesis on the compensation of linear effects.

2.2.1 Dispersion compensation & polarization de-multiplexing

Dispersive effects can be mitigated by time-domain or frequency-domain equalization. In a single-carrier context, a famous time-domain equalization technique is the concatenation of a finite impulse response (FIR) filter for the quasi-deterministic polarization independent CD with a Constant Modulus Algorithm (CMA) for PMD, residual CD and polarization de-multiplexing. Using Fourier transforms, the transfer function of CD in (1.12) can be inverted to obtain a time-domain CD compensating filter having the following tap weights [25]:

$$g(m) = \sqrt{\frac{icT_S^2}{D\lambda^2 z}} \exp\left(-i \frac{\pi cT_S^2}{4D\lambda^2 z} m^2\right) \quad (2.4)$$

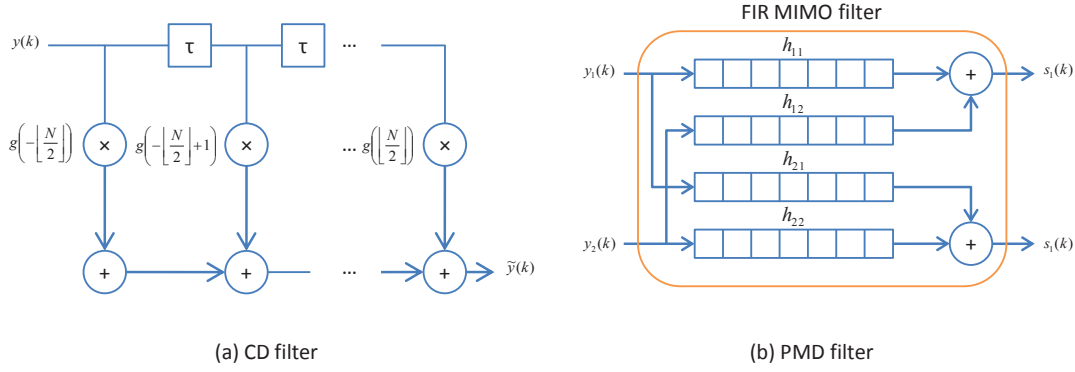


Figure 2.2: Time-domain equalization: (a) FIR filter structure for compensating for chromatic dispersion, (b) Butterfly structure to compensate for polarization-mode-dispersion and polarization crosstalk.

with $[-N/2] \leq m \leq [N/2]$ and $N = \lfloor \frac{|D|\lambda^2 z}{2cT_s^2} \rfloor$ being the total number of filter taps. The same filter, shown in Fig. 2.2(a), is applied on $\mathbf{y}_x$ and $\mathbf{y}_y$ and can compensate for all CD or residual CD if in-line compensating fibers (DCFs) with a negative dispersion coefficient are used. However, DCFs usually have a large non-linear coefficient which makes the all-digital CD compensation a better solution.

As for the inter-symbol interference and polarization crosstalk resulting from PMD and birefringence in PDM systems, a MIMO approach implemented through a butterfly FIR filter, as shown in Fig. 2.2(b) is employed to remove interference and retrieve the multiplexed signals. Indeed, the received signal on each polarization can be expressed as:

$$y_i(k) = \sum_{j=1}^2 \sum_{l=0}^{L-1} h_{i,j}(l;k) s_j(k-l) + n_i(k), \quad i = \{1, 2\} \quad (2.5)$$

where $i, j = 1, 2$ denote the x - and the y - polarizations respectively, L the channel memory corresponding to the maximum DGD in number of samples T_e , l an index related to the delay variable and k related to the time variable. The channel between the i^{th} input and the j^{th} output consists of a vector $\mathbf{h}_{ij}(k) = [h_{ij}(1,k) \dots h_{ij}(l,k) \dots h_{ij}(L,k)]$ with L taps and adaptive weights depicting the time dependence of the channel.

At the receiver, the signals $\mathbf{y}_i$ are sampled, for instance at twice its maximum frequency. The MIMO equalizer consists of four FIR filters $\mathbf{g}_{ij}(k) = [g_{ij}(1,k) \dots g_{ij}(l,k) \dots g_{ij}(L_f,k)]$ of L_f taps. The estimate of the symbol transmitted at the j^{th} polarization at time instant k can be written in function of the filter taps and the received symbols as:

$$\hat{s}_j(k) = \sum_{i=1}^2 \left[\sum_{l=0}^{L_f-1} g_{i,j}(l;k) y_i(k-l) \right], \quad j = \{1, 2\} \quad (2.6)$$

The equalizer tries to adapt the coefficients of the four filters to the inverse of the Jones matrix of the optical channel. If the number of filter taps is large enough to include the maximum delay induced by dispersive effects $L_f \geq L$, inter-symbol interference can be completely eliminated.

The separation of CD compensation from the adaptive equalization of PMD and crosstalk is desirable since it leads to shorter adaptive filters since PMD accumulates much slower than CD [25]. Different update algorithms exist depending on the chosen modulation formats. The most widely used algorithm for PDM-QPSK systems is the CMA which is a blind adaptive equalizer that exploits the fact that the symbols lie on the same circle (i.e. have a constant modulus). It is based on a least-mean square algorithm that tries to minimize the difference between the filter coefficients and the inverse channel matrix [114]. Polarization fluctuations up to 50 kHz can be tracked using this algorithm [115]. An improvement can be brought to this standard CMA equalizer through switching to a decision directed mode after the convergence of the CMA [25]. The standard CMA can also be adjusted for multi-radius constellations such as 16-QAM [116].

On the other hand, frequency-domain equalization (FDE) can be done both with single-carrier or multi-carrier schemes. The main advantage of FDE over time-domain equalization is the size of the equalizer that does not linearly scale with the channel memory. Indeed, both CD and PMD can be compensated for using the inverse of the 2×2 frequency-domain transfer matrix of the channel $\mathbf{H}$. In this case, an FFT/iFFT block is needed to perform the time-frequency and frequency-time conversions and the complexity of the equalization block is determined by the size of the FFT/iFFT multiplied by the complexity of the inversion of one 2×2 matrix per frequency component, whereas the complexity of a time-domain equalization (TDE) exponentially grows with the amount of dispersion in the link (more accumulated dispersion induces larger delays, hence larger filters with more taps L_f). The inversion of the channel matrix implies its knowledge at the receiver side. In practice, the channel matrix is periodically estimated through a training sequence known at the receiver (usually on a millisecond scale, following the temporal variations of PMD) which gives a better estimation of the channel and avoids the convergence problem of blind equalization techniques such as CMA [25].

2.2.1.1 Scale-up for MDM systems

For an MDM system with $2M$ modes ($2M$ includes spatial and polarization modes), the 2×2 equalizers used for PDM systems have to be scaled up to $2M \times 2M$ MIMO systems requiring $4M^2$ adaptive FIR filters to remove modal crosstalk. Moreover, to completely compensate DMGD, the length of the filters should be larger than the maximum differential modal delay which can easily accumulate and reach larger delays than the ones induced by PMD. An MDM system can be represented as a frequency-selective and time-varying MIMO channel. By expanding (2.5) to

an $2M \times 2M$ system, the received signal on each mode is given by:

$$y_i(k) = \sum_{j=1}^{2M} \sum_{l=1}^L h_{i,j}(l; k) s_j(k-l) + n_i(k), \quad i = \{1, \dots, 2M\} \quad (2.7)$$

We have seen in section 1.4.4 that among the best reported DMGD in few-mode-fibers were 6 ps/km for a 3-mode fiber and 40 ps/km for a 6-mode fiber compared to SMFs with a PMD coefficient as low as 0.05 ps/ $\sqrt{\text{km}}$. We have also mentioned that strong modal coupling or DMGD cancellation can slow down the accumulation of DMGD in the optical transmission system. Despite all this, maintaining low DMGD values at the end of long-distance links is challenging. In [104], the authors point out that the computational complexity of TDE methods increases exponentially with the total DMGD of the link making TDE unsuitable for long-distance MDM transmissions. Single-carrier FDE has been proposed in [117, 118] with a complexity scaling in a logarithmic fashion with the total DMGD, whereas multi-carrier FDE techniques, in particular OFDM, have been proven to be capable of eliminating all ISI induced by DMGD with the lowest complexity [104]. However, the reach of MDM-OFDM systems, or equivalently the maximum allowable DMGD, is limited by the length of the cyclic prefix of the OFDM symbol. More details on OFDM modulation will be given later in the chapter.

2.2.2 Frequency offset compensation & carrier phase estimation

The local oscillator at the receiver side may exhibit a frequency offset from the incident optical carrier (mostly in intradyne coherent detection where the LO can freely run within 1 GHz of the optical carrier) creating a progressive phase shift $\Delta\Phi$ in the received signals. The phase of any of the two time-domain equalized signals $\hat{s}_{j=\{1,2\}}(k)$ in a PDM system (respectively $2M$ in an MDM system) is then given by:

$$\arg(\hat{s}_j(k)) = \Phi_S(k) + \Phi_0(k) + k\Delta\Phi \quad (2.8)$$

where Φ_S is the modulation phase and Φ_0 a random phase offset. The offset can be estimated in the frequency or in the time domain, in order to correct the induced phase shift. Both ways rely on a 4^{th} power operation to remove the modulation. In the frequency domain method, the 4^{th} power of the equalized vector $\hat{\mathbf{s}}_j = [\hat{s}_j(1), \dots, \hat{s}_j(K)]$ is taken, followed by an FFT operation and a search for the maximum of the resultant signal that will be located at the offset frequency as shown below:

$$\begin{aligned} C(f) &= \text{FFT}(\hat{\mathbf{s}}_j^4) \\ \widehat{\Delta f} &= \arg\max(C(f)) \\ \widehat{\Delta\Phi} &= \frac{2\pi\widehat{\Delta f}}{T_e} \\ \hat{s}'_j(k) &= \hat{s}_j(k) \exp(-ik\widehat{\Delta\Phi}) \end{aligned} \quad (2.9)$$

where FFT is the fast Fourier transform operation, T_e the sampling period and $(\cdot)^4$ is the component-wise fourth-power operation. This method is accurate while being computationally intensive. It may be used for short signals for simulations and experiments. In a real system, a more suitable time-domain method can be implemented. To estimate the phase shift $\Delta\Phi$, the first two phases in (2.8) have to be eliminated, and the phase shift averaged over a block of N symbols [119] such that:

$$\widehat{\Delta\Phi} = \frac{1}{4} \arg \left(\frac{1}{N} \sum_{m=-N/2+1}^{N/2} (\hat{s}_j(k+m) \hat{s}_j^*(k+m-1))^4 \right) \quad (2.10)$$

where $(\cdot)^*$ is the conjugate operation.

When the frequency offset is removed, the carrier phase offset has to be estimated and eliminated before demodulating the symbols. The random phase offset Φ_0 results from phase noise generated by both the LO and transmitter lasers. This phase noise is common for all the multiplexed signals (polarizations and modes) and is generally modeled as a Wiener process of variance $\sigma^2 = 2\pi\Delta\zeta T_S$ with $\Delta\zeta$ representing the laser linewidth. Typical linewidth values of external cavity lasers (ECL) are around 100 kHz and can go up to a few MHz for distributed feedback (DFB) lasers. A commonly used carrier phase recovery method for QPSK signals is a feed-forward phase estimation known as the Viterbi and Viterbi algorithm [11, Chap.5] that also relies on a fourth power operation followed by an averaging given by:

$$\widehat{\Phi}_0(k) = \frac{1}{4} \arg \left(\frac{1}{N} \sum_{m=-N/2+1}^{N/2} \hat{s}_j(k+m)^4 \right) \quad (2.11)$$

The estimated phase offset can then be applied to the averaging window. The optimum size of this window depends on the linewidth of the lasers and the symbol rate. A lower symbol rate increases the laser phase noise variance and leads to larger phase walk-offs. Also note that the fourth power operation in (2.11) creates a phase ambiguity of $\pm a\pi/2$, ($a \in \mathbb{N}^*$) that must be combated by unwrapping the estimated phase to avoid cycle-slipping. Differential encoding was adopted in optical communication systems to mitigate detection errors caused by cycle-slipping. Moreover, this carrier recovery method can only be generalized to PSK modulations and other algorithms should be used for multi-level modulations such as q-QAM [116].

Once the carrier phase is recovered, the signals are ready for demodulation and FEC decoding. As discussed above in section 2.1, a hard-decision approach can be realized through a partitioning of the constellation into decision regions followed by thresholding and two-level quantization (0s and 1s), or else we can feed the FEC decoder with more finely quantified information called soft values, usually consisting of log likelihood ratios (LLR) representing the probability of a bit to be a 0 or a 1 [11, Chap.8].

2.2.3 Transmitter-side DSP

It is worthy to note that basic DSP can also be applied on the transmitter (TX) side, for instance through the insertion of DACs in order to implement less-complex optical transmitters with multi-level modulation formats. Transmit-side DSP can also be realized through spectrum pulse shaping to enable the transmission of super-channels made of several closely packed carriers.

More sophisticated DSP at the transmitter can also be envisaged such as pre-distortion of the signal to combat some channel effects. However, this will be the most effective when the channel properties are known at the transmitter – in this case it is referred to as *a priori* information at TX. Knowing that the optical channel exhibits random effects such as PMD and non-linearity, a feedback signal from the receiver would be required to perform the DSP. Hence, for long distances covered by regional and long-haul optical links, any transmitter-based DSP that needs a priori information will be impractical.

2.3 OFDM: more than a modulation format

Among multi-carrier communication techniques, OFDM is the most popular format and has been adopted in wireless multi-antenna communication standards such as IEEE 802.11 (Wi-Fi) to combat multi-path interference and better suit frequency-selective channels [120]. OFDM also paved the way to the implementation of new coding schemes for MIMO channels such as ST codes that we will introduce in the next section. The management of interference and delay by the OFDM format allows for a low-complexity single-tap MIMO equalization at the receiver, which enables the use of sophisticated MIMO schemes to compensate for some other channel impairments.

With the adoption of coherent detection in optical communications and recent advancements in high-speed CMOS electronic circuits, OFDM has emerged as an attractive multi-carrier solution for 100 Gb/s to 1 Tb/s transmission systems. We recall in this section the architecture of an OFDM-based transmission system and the associated DSP as well as the advantages that OFDM offers for optical MIMO systems.

2.3.1 Principle

The idea of OFDM is to transmit data over a large number of low-rate orthogonal subcarriers, or parallel streams, such that each low-rate subcarrier observes a channel with a flat-frequency response as shown in Fig. 2.3, resulting in a longer symbol duration and a better tolerance to ISI. In the time domain, summing the signals over all the subcarriers gives us a complex-valued analog signal with Gaussian distributed component. Furthermore, a guard interval is inserted between the successive OFDM symbols and an appropriate windowing of the received symbols is applied in order to retrieve the useful data that remains uncontaminated by ISI as seen in Fig. 2.3.

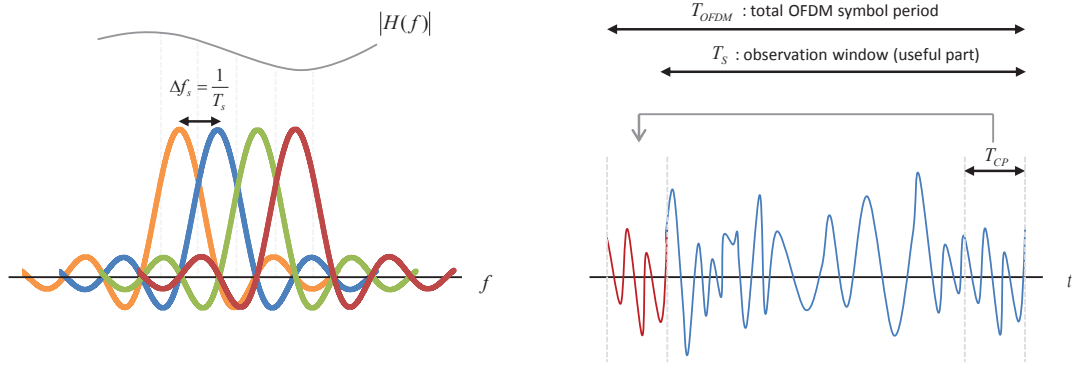


Figure 2.3: OFDM signal: on the left, the OFDM spectrum and a random channel frequency response; on the right, a time-domain OFDM symbol with a cyclic prefix.

Let N_{sc} be the number of OFDM subcarriers. N_{sc} subcarriers with overlapping spectra are chosen and their orthogonality is guaranteed if they are equally spaced at multiples of the inverse of the symbol rate $\Delta f_s = 1/T_s$, as can be seen in Fig. 2.3. This subcarrier packing increases the spectral efficiency of the scheme and avoids the need for narrow filters for each subcarrier. However, due to multi-path interference or dispersion in the channel, the subcarriers may lose their orthogonality. To avoid this from happening, a guard interval, also called a cyclic prefix (CP) which is a copy of a small part of the end of the OFDM symbol, is inserted at the beginning of each symbol to absorb interference and preserve at the same time the orthogonality of the subcarriers [120]. Other channel effects can harm the subcarrier orthogonality such as frequency offsets and carrier phase noise at the LO. Dedicated algorithms compensating both effects will be discussed in the next sections.

2.3.2 OFDM for optical communications

Multi-carrier transmission schemes already exist in optical fiber transmission links through WDM for instance. However, the use of multi-carrier modulations within each optical carrier is a new trend that was mainly pioneered by coherent detection and the need to achieve higher data rates exceeding 40 Gb/s. These new high-capacity systems are more vulnerable to chromatic and polarization mode dispersion because the induced ISI becomes stronger. Although single-carrier based techniques were quickly enhanced to address this increased ISI and made their way into the latest 100 Gb/s transceivers, the intrinsic strengths of OFDM including ISI elimination, channel estimation and high-spectral efficiency have been also recognized by the researches in the optical communication systems [103, 121]. Moreover, OFDM is compatible with optical MIMO systems, for instance with PDM schemes where two independent OFDM signals can be generated and combined in the fiber by a polarization beam combiner. In an optical PDM-OFDM system, the

cyclic prefix should be larger than the total amount of dispersion in order to completely eliminate ISI [103]. This condition is given by:

$$T_{CP} \geq \frac{c}{f_0^2} |D_{acc}| N_{sc} \Delta f_S + \text{DGD}_{max} \quad (2.12)$$

where T_{CP} is the cyclic prefix duration, D_{acc} the total accumulated dispersion in ps/nm and DGD_{max} the maximum budgeted differential group delay that is usually chosen as three times the average DGD in the link. After adding the CP, the total OFDM symbol duration is $T_{OFDM} = T_S + T_{CP}$.

OFDM can be implemented in several ways depending on how we generate the multi-carrier spectrum. A first category relies on all-optically generated or demodulated orthogonal subcarriers in which tens of Tb/s total data rate can be achieved thanks to a large available bandwidth given by the product of the DAC/ADC sampling rate times the number of optical subcarriers [122]. Optical signal-processing devices are thus needed to perform the FFT/iFFT operations. Another category consists in modulating the electrical OFDM spectrum on an optical carrier. This can be done through a simple intensity modulation along with direct detection at the receiver, or through a complex (intensity and phase) modulation along with coherent detection. The former is referred to as direct detection optical (DDO-) OFDM and is more attractive for short-reach access network links that require a cost-effective solution [123]. The latter is called coherent optical (CO-) OFDM and is more suitable for high-capacity long-haul optical links where spectral efficiency and reach are the most important parameters [103]. In these systems, the OFDM waveform is electronically generated and demodulated through FFT/iFFT operations and DAC/ADC converters and the signal bandwidth is limited by the sampling rate of the converters. To overcome this limitation, a multi-band CO-OFDM approach was suggested and consists in splitting the total bitrate into several OFDM sub-bands generated by different DACs and transmitted at neighboring wavelengths (e.g. 100 Gb/s divided into 4 OFDM sub-bands of 25 Gb/s each [124]). In the following, we will detail the transmitter and the receiver sides of a PDM CO-OFDM system that offers the ultimate performance in receiver sensitivity, spectral efficiency and robustness against polarization effects [125].

2.3.2.1 PDM CO-OFDM transmitter

OFDM modulation can be implemented by an inverse Fast Fourier transform (iFFT). A block of N_{sc} modulated symbols $\mathbf{s} = [s_1, \dots, s_{N_{sc}}]$ is fed to an iFFT block to convert it to the time domain into a vector $\tilde{\mathbf{x}} = [\tilde{x}_1, \dots, \tilde{x}_{N_{sc}}]$ such that for $m = \{1, \dots, N_{sc}\}$, we get:

$$\tilde{x}_m = \frac{1}{\sqrt{N_{sc}}} \sum_{k=1}^{N_{sc}} s_k \exp\left(\frac{i2\pi km}{N_{sc}}\right) \quad (2.13)$$

where s_k is the complex data symbol transmitted by the k^{th} subcarrier. The cyclic prefix of length $\Delta_G = \lceil T_{CP}/T_e \rceil$ where T_e is the sampling period, is then added by repeating a part of the last components of the output vector at its start, thus obtaining a vector $\tilde{\mathbf{x}} = [\tilde{x}_{N_{sc}-\Delta_G+1} \dots \tilde{x}_{N_{sc}}, \tilde{x}_1, \dots, \tilde{x}_{N_{sc}}]$. To completely avoid ISI, the cyclic prefix must be larger than the maximum delay spread caused by the channel. Apart from the useful data, the subcarriers can carry a synchronization sequence, a training sequence and pilot symbols to perform time and frequency synchronization, channel estimation and carrier phase recovery at the receiver. These extra symbols along with the cyclic prefix represent an overhead that must be accounted for in the total bitrate.

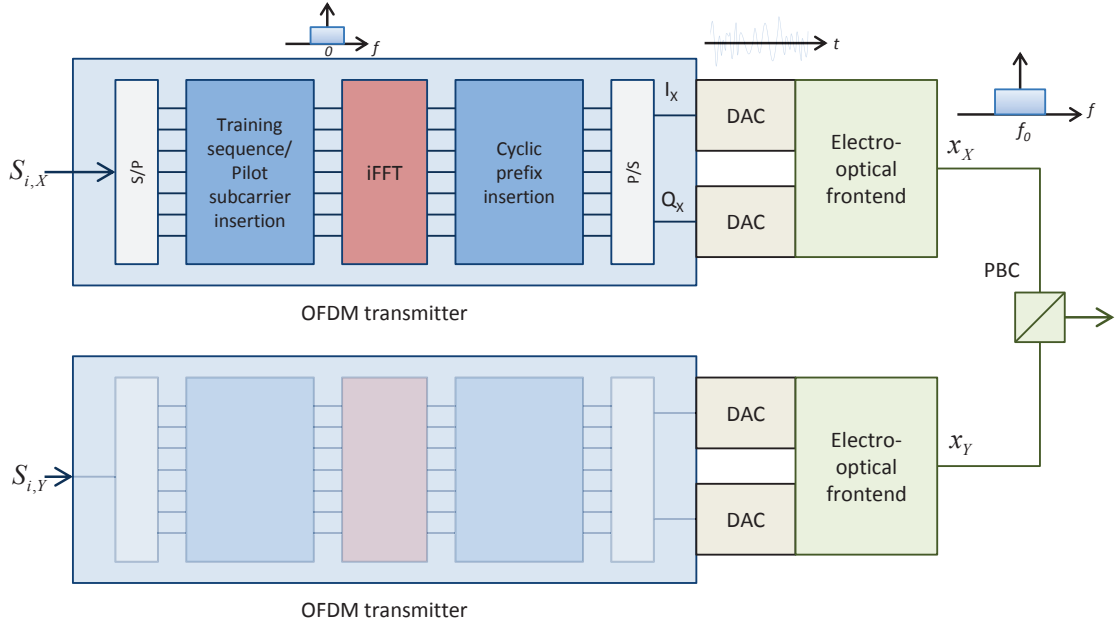


Figure 2.4: Polarization division multiplexed OFDM transmitter with the main OFDM blocks.

Fig. 2.4 shows a PDM-OFDM transmitter. Two independent OFDM signals are generated by multiplexing data symbols on subcarriers through the iFFT block. The data symbols can be represented as a $2 \times N_{sc}$ matrix $\mathbf{S} = [\mathbf{s}_1 \dots \mathbf{s}_{N_{sc}}]$ where $\mathbf{s}_k = [s_k^{(X)}, s_k^{(Y)}]^T$ are the two data symbols transmitted on two polarizations by the k^{th} subcarrier. Then, at the output of the iFFT block, a cyclic prefix is inserted and the sampled digital signals are converted to analog signals by DACs, then filtered by low-pass filters to eliminate the aliasing replicas of the baseband OFDM spectrum. After, they are up-converted to the optical domain using an MZM I/Q modulator as the one shown in Fig 1.14 for PDM-QPSK with the binary signals replaced by the OFDM signals. Given their sinusoidal transfer functions in Fig. 1.4, the MZMs are biased at their null points (trough) to achieve a linear conversion between the baseband electrical OFDM signal and the optical field. Other RF-to-optical conversions may be used such as up-converting the baseband signal to an intermediate RF frequency that modulates a single MZM followed by an optical band-pass filter to eliminate the undesired image side band [43].

2.3.2.2 PDM CO-OFDM receiver

At the receiver, a polarization beam splitter separates the incoming signal by projecting it on two orthogonal polarization states. The signals are then down-converted to the optical domain through a dual-polarization coherent receiver as the one in Fig 1.14. After analog-to-digital conversion of the signals, any frequency offset between the LO and the optical carrier, as well as any sampling frequency offset, should be compensated for. Then, a time synchronization is required in order to determine the beginning of the OFDM frames. Once the data frames are located, OFDM demodulation can be carried through the removal of the cyclic prefix and an FFT block to convert the signal back to the frequency domain. The presence of the cyclic prefix and the compensation of the frequency offset maintain the orthogonality of the subcarriers, hence the FFT output $\mathbf{y}_k = [y_k^{(X)}, y_k^{(Y)}]^T$ linearly depends only on the iFFT input $\mathbf{s}_k$ at each subcarrier $k = \{1 \dots N_{sc}\}$ [121] such that:

$$\mathbf{y}_k = \exp(i\Phi_0) \mathbf{H}_k \mathbf{s}_k + \mathbf{n}_k \quad 1 \leq k \leq N_{sc} \quad (2.14)$$

where $\mathbf{H}_k$ is the 2×2 frequency-domain transfer matrix of the channel for the k^{th} subcarrier, Φ_0 is a phase noise term caused by the lasers at the transmitter and the receiver, and $\mathbf{n}_k$ a noise vector. In (2.14), we have neglected any fiber non-linearity. The channel matrix is frequency dependent due to the presence of CD and PMD in the optical link. Both the properties of the optical channel and the sources of phase noise must be taken into account when designing the OFDM system. Indeed, while it is preferable to have a large number of subcarriers in order to limit the overhead $\Delta_G / (N_{sc} + \Delta_G)$ introduced by the CP, this would lead to narrower subcarriers or equivalently a longer symbol duration and hence less tolerance to phase noise [126]. At the same time, if a small number of subcarriers is considered, the frequency-domain transfer matrix $\mathbf{H}_k$ may no longer be constant over the subcarrier bandwidth (mainly due to the frequency dependence of CD and PMD) [121].

The optical channel also exhibits rather slow temporal variations and may be constant over an OFDM frame. Indeed, if we consider a 100 Gb/s OFDM transmission with $N_{sc} = 256$ subcarriers, the duration of a single OFDM symbol is roughly $256/10^{10} = 2.56$ ns (neglecting CP which is usually kept as low as possible to limit the overhead) which is far lower than the dynamics of PMD in installed fibers that change over a millisecond scale, and also lower than the fastest polarization rotations encountered in lab experiments (microsecond scale). Therefore, after removing the phase error Φ_0 , we can extend (2.14) that corresponds to the k^{th} subcarrier of a single OFDM symbol to a larger OFDM frame consisting of N_{sym} symbols and obtain:

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{S}_k + \mathbf{N}_k \quad 1 \leq k \leq N_{sc} \quad (2.15)$$

2. ADVANCED SIGNAL PROCESSING AND CODING TECHNIQUES FOR OPTICAL COMMUNICATIONS

where $\mathbf{Y}_k$ (resp. $\mathbf{S}_k$) is the $2 \times N_{sym}$ received (resp. transmitted) symbols over the k^{th} subcarrier and in N_{sym} consecutive OFDM symbols and $\mathbf{N}_k$ is a $2 \times N_{sym}$ matrix representing the additive noise. A joint MIMO processing of the two polarization-multiplexed OFDM signals must now be carried in order to retrieve the data symbols and demodulate them. To do so, the channel matrix $\mathbf{H}_k$ must be first estimated, then inverted. The PDM CO-OFDM digital receiver architecture is given in Fig. 2.5 and its main blocks are detailed below.

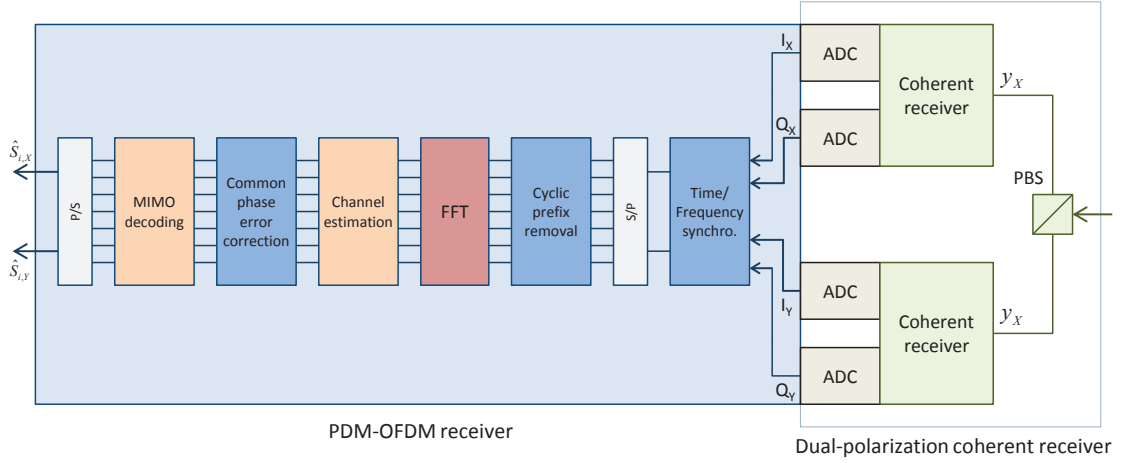


Figure 2.5: Coherent polarization division multiplexed OFDM receiver with the main OFDM blocks.

Time and frequency synchronization

Accurate temporal location of the start of an OFDM frame as well as accurate location of the OFDM subcarriers are crucial as otherwise the orthogonality of OFDM subcarriers can be broken leading to both ISI and inter-carrier interference (ICI) [127]. Several algorithms were suggested to perform these synchronizations. They all rely on special OFDM training symbols such as an OFDM symbol with identical halves in the time domain or by X-fold repetition of a pattern with a possible alternation of signs. These symbols are then detected through an auto-correlation operation over a sliding window followed by a search for the peak of the auto-correlation metric to determine the start of an OFDM frame. The most famous estimator of this form was suggested by Schmidl and Cox [128] where two training symbols with two identical halves are used for frequency and time synchronization. However, this algorithm has a limited precision because its metric exhibits a plateau rather than a precise peak. More robust algorithms were later suggested by Minn and Barghava [129] or Shi and Serpedin [130]. Although these algorithms were initially designed for single-input single-output (SISO) systems, extended versions for MIMO schemes have been thoroughly investigated for wireless channels [120].

As for frequency offset resulting from a carrier frequency offset (CFO) and a sampling frequency offset (SFO), it can be normalized with respect to the subcarrier spacing to get $\varepsilon = \Delta f / \Delta f_s$. This normalized offset can be large and may vary with time due to the wavelength instability of the lasers. The above-mentioned time-synchronization algorithms are also capable of estimating frequency offsets using the repetitive structure of their training symbols. They operate in two steps by computing the integer and fractional parts of ε separately, then correcting the progressive phase shift induced by this offset in the time-domain OFDM signals. However, the range of the frequency offset estimate of these algorithms depends on the length and the structure of their training sequences and is hence limited. Therefore, should the real offset be greater than the range of the chosen algorithm, a frequency-domain CFO correction step can be appended to remove the residual frequency offset [120, 131].

Channel estimation

The channel $\mathbf{H}_k$ can be estimated using training symbols $\mathbf{TS}_k$ over the subcarriers. Having a 2×2 channel matrix, the training symbols have to be sent on both polarizations and during two time slots at least in order to estimate the channel coefficients. A simple and popular implementation of a training sequence was suggested in [132] where known symbols are sent in a first OFDM symbol on the x -polarization while no information (zeros) are sent over the y -polarization, then the opposite is done for the second symbol (known symbols over the y -pol. and zeros over the x -pol.). At the k^{th} subcarrier, the training sequence $\mathbf{TS}_k$ can be written as the 2×2 matrix:

$$\mathbf{TS}_k = \begin{bmatrix} s_k^{(X)}(t_1) & 0 \\ 0 & s_k^{(Y)}(t_2) \end{bmatrix} \quad (2.16)$$

where $s_k^{(X)}(t_1)$ belongs to the OFDM training symbol sent on the x -pol. during the time slot t_1 , and $s_k^{(Y)}(t_2)$ belongs to the second OFDM training symbol sent on the y -pol. during the time slot $t_2 = t_1 + T_{OFDM}$. Several estimation algorithms exist to retrieve $\mathbf{H}_k$, among which we choose the one based on the least square estimation (LSE) criterion [132] that is popular for its simple implementation. From (2.15) and (2.16), the receiver can estimate the transfer matrix at each subcarrier $\hat{\mathbf{H}}_k$ by applying the following:

$$\hat{\mathbf{H}}_k = \mathbf{Y}_k \left(\mathbf{S}_k^\dagger (\mathbf{S}_k \mathbf{S}_k^\dagger)^{-1} \right) \quad (2.17)$$

where the training sequence $\mathbf{S}_k$ is known at the receiver. Using the training sequence structure in (2.16), the matrix products in (2.17) reduce into simple scalar products. In order to attenuate the estimation noise, several OFDM symbols can be sent for training and the estimated transfer

matrices are obtained by time-domain averaging. The training sequence can be sent in the preamble of each OFDM frame just after the symbols used for synchronization, in order to periodically estimate the channel. An example of a structured OFDM frame is given in Fig. 2.6. The typical duration of OFDM frames is few milliseconds depending on the dynamics of the channel. In each frame, 10 to 50 OFDM symbols are used for channel estimation [132].

Finally, we note that the sequence in (2.16) is not unique nor optimal, but simple to implement, and that other structures of training sequences for MIMO-OFDM systems were proposed in literature with a better power balance between the two data-carrying channels [120]. Moreover, in order to limit the overhead brought by the preamble, frequency-domain averaging and interpolation techniques were suggested in [133] that permit to reduce the number of OFDM training symbols.

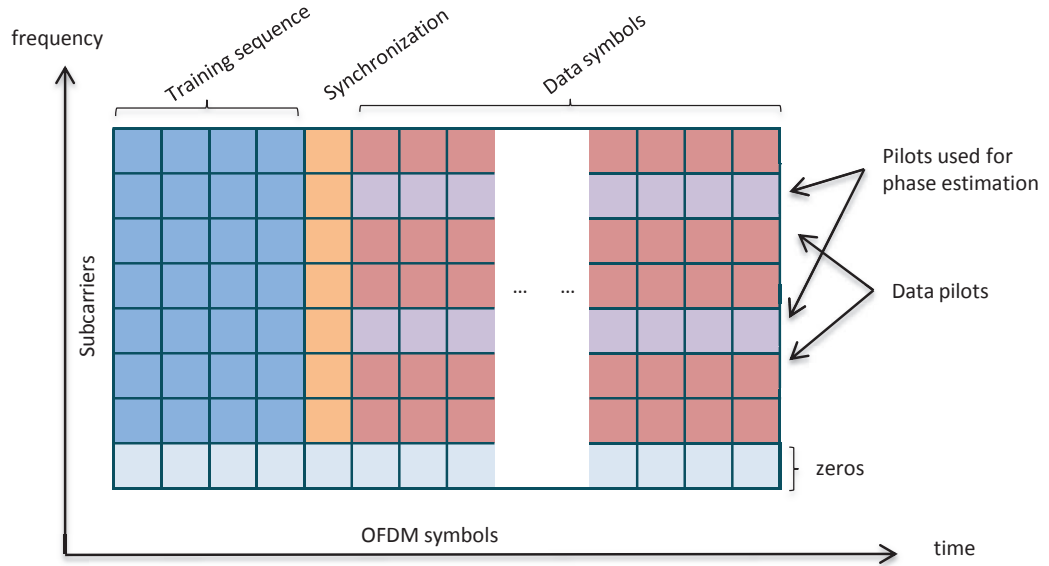


Figure 2.6: OFDM frame design on the frequency-time grid showing the locations of training and synchronization symbols.

Phase noise compensation

CO-OFDM implementation requires high-quality lasers with 100 kHz linewidth or even lower in order to slow down the evolution of random phase shifts introduced by the lasers at the TX and RX. Given that the duration of an OFDM symbol is far larger than a single-carrier transmitted symbol, the phase noise can change within a single symbol if lasers with a large linewidth were used, making it impossible to estimate and correct the phase error. Provided that the phase noise trajectory slowly evolves along the OFDM symbol duration, the phase noise has two main impacts on OFDM transmissions [126]: a phase noise error Φ_0 that is common to all subcarriers of a

single OFDM symbol, called common phase error (CPE) and an ICI term modeled as additive white Gaussian noise.

At the receiver, two main methods were proposed in order to compensate for phase noise. The first one, already used in wireless OFDM systems, uses N_p pilot tones in each OFDM symbol (known data at certain subcarriers shown in Fig. 2.6) in order to estimate $\widehat{\Phi}_0$ and correct it without addressing ICI [126]. Having estimated the channel and knowing the pilot symbols at the receiver, the receiver equalizes the signals over the N_p subcarriers and deduces a phase noise estimate by averaging. The second method, suggested by Jansen in [134], consists in inserting an RF pilot tone in the middle of the OFDM spectrum. This tone experiences the same phase noise as the OFDM signal and will be filtered out at the receiver, conjugated then multiplied with the OFDM signal in order to remove CPE and reduce ICI.

Symbol detection

Given that a suitable cyclic prefix absorbed all the interference, a single-tap MIMO equalizer is enough to de-multiplex the polarizations instead of the multi-tap equalizer shown in Fig. 2.2. The least-complex practical implementation of a MIMO equalizer is the zero-forcing (ZF) method which consists in removing the polarization crosstalk using the inverse of the estimated channel matrix at each subcarrier such that:

$$\widehat{\mathbf{S}}_k = \mathbf{W}_k \mathbf{Y}_k \quad \text{with} \quad \mathbf{W}_k = \left(\widehat{\mathbf{H}}_k^\dagger \widehat{\mathbf{H}}_k \right)^{-1} \widehat{\mathbf{H}}_k^\dagger \quad (2.18)$$

where $\mathbf{W}_k$ is an equalization matrix selected as the pseudo-inverse of the channel matrix that coincides with the inverse for square MIMO channels. We shall see in the next section, where we look at MIMO coding and decoding, that ZF equalization has a very poor performance over channels with random fading such as wireless channels or non-unitary optical channels because the channel inversion can lead to an important amplification of noise.

2.3.3 Capacity of linear MIMO-OFDM channels

Later in the thesis, we will be studying capacity-limiting effects in the optical MIMO channels and measuring their capacities. We recall in this section the definitions of the capacities of linear MIMO channels. The channel capacity C , also known as Shannon capacity, is defined as the upper bound of the achievable transmission rates R for an arbitrary small error probability. C is usually expressed in bits/s/Hz. Two kinds of capacities can be distinguished depending on the dynamics of the channel. By comparing the symbol duration T_S to the coherence time of the channel T_c , a parameter that indicates how fast a channel varies in time, we can distinguish fast-fading channels where $T_S \sim T_c$ from slow-fading channel where $T_c \gg T_S$.

For fast-fading channels varying at each symbol time T_S , we define an “ergodic capacity” that is the mean capacity averaged over all the channel realizations. In MIMO-OFDM systems, the total capacity is the sum of the capacities of each subcarrier that observes an $N \times N$ channel with a flat frequency response given by:

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{S}_k + \mathbf{N}_k \quad (2.19)$$

where $\mathbf{N}_k$ represents a circularly symmetric additive white Gaussian noise. Each of its components is a complex zero-mean Gaussian variable of variance N_0 per real dimension. With channel knowledge at the receiver, the ergodic capacity of each subcarrier is equal to [135]:

$$C_k = \mathbb{E}_{\mathbf{H}_k} \left\{ \log_2 \left(\det \left(\mathbf{I} + \frac{\rho}{N} \mathbf{H}_k \mathbf{H}_k^\dagger \right) \right) \right\} = \sum_{i=1}^N \mathbb{E}_{\mathbf{H}_k} \left\{ \log_2 \left(1 + \frac{\rho}{N} \lambda_i^2 \right) \right\} \quad (2.20)$$

where λ_i are the singular values of $\mathbf{H}_k$ assumed to be full-rank and let $\rho = \frac{E_{tot}}{2N_0}$ be the average signal-to-noise ratio (SNR) at the receiver on the k^{th} subcarrier. E_{tot} is the average total energy enforced at the transmitter. The capacity expression in (2.20) neglects the non-linear effects in the optical channel which will result in a capacity decrease at high SNR, or equivalently at high injected signal power, instead of a monotonously increasing capacity as a function of the SNR [6].

The second capacity definition is used for slowly-varying channels (for instance block-fading channels where T_c equals the duration of a block of T symbols). In this case, the capacity is seen as a random variable and a useful performance metric is the probability that the capacity falls below a certain value called the outage capacity C_p for a specified percentage of channel realizations. Assuming that $\mathbf{H}_k$ is quasi-static for a block of data but changes from a block to another, the probability:

$$\Pr(C_k \leq C_p) = P_{out} \quad (2.21)$$

is called the outage probability and the subscript p in C_p denotes P_{out} . For each channel realization, C_k is computed as in (2.20) but without the averaging operation. For any given SNR, the probability density function (PDF) of C_k and from it, the cumulative distribution function (CDF) of C_k can be drawn and the outage capacity at some particular P_{out} is easily determined from this CDF. (2.21) can be interpreted as a failure of transmission at a rate C_p with probability P_{out} .

In Information Theory and Communication Theory communities, the SNR is the most commonly used parameter against which the capacity or the BER performance of a given linear MIMO channel is measured. We can distinguish a symbol-based SNR or simply SNR and an information-bit based SNR denoted as SNR_{bit} related by:

$$\text{SNR} = \frac{E_S}{2N_0} = \frac{\rho}{N} = \frac{k_c r_{FEC} E_b}{2N_0} = k_c r_{FEC} \text{SNR}_{bit} \quad (2.22)$$

where E_S is the average symbol energy in $\mathbf{S}_k$, k_c the spectral efficiency of the modulation format, r_{FEC} is the FEC overhead and $2N_0$ is the total spectral density of the noise over one of the N dimensions of the MIMO channel. In optical communications, the OSNR defined in (1.10) is rather used and can be related to the SNR by [6]:

$$\text{OSNR} = \frac{R_b}{2B_{ref}} \text{SNR} \quad (2.23)$$

where R_b is the information bitrate and B_{ref} is the reference spectral bandwidth within which the ASE noise is observed (usually corresponds to 0.1nm). (2.23) is obtained by taking $2N_0 = N_{spans}N_{ASE}$.

2.3.4 Optical OFDM or single-carrier for MIMO schemes?

The key advantages of CO-OFDM over SC in optical MIMO systems are the low-complexity single-tap equalization and the channel estimation that provides important information about the impairments of the optical channel and offers the possibility to digitally compensate them. Compensation of CD and PMD as well as polarization de-multiplexing were already demonstrated in many CO-OFDM experiments [43, 103].

However, more can be done with the knowledge of the channel at the receiver such as the application of MIMO codes to mitigate the non-unitary impairments of the MIMO channel as we shall see later in this chapter. These MIMO codes correlate the data symbols among the degrees of freedom of the MIMO channel in order to protect the transmitted information, and hence necessitate a more sophisticated decoding. The low-complexity frequency-domain equalization of MIMO-OFDM schemes paves the way for the implementation of MIMO codes with a reasonable decoding complexity. Indeed, while the complexity of ISI equalization in SC schemes grows exponentially with the number of taps at the RX, the complexity of OFDM is shared between the TX and the RX through the iFFT/FFT operations. The complexity of these operations is insensitive to the number of taps and scales moderately with the number of subcarriers N_{sc} .

OFDM and MIMO form a robust scheme for very high spectral-efficiency multiplexed optical systems because the linear interference from neighboring channels (polarizations or modes) at the same frequency can be reversed by a matrix inversion operation or more sophisticated decoders using a single-tap filter that replaces the multi-tap time-domain filter of current SC-based schemes. In particular, few-mode-fiber (FMF)-based transmissions associated with a MIMO-OFDM scheme has shown to be a promising technology to attain capacities higher than 100 Tb/s with the lowest complexity compared to a large set of TDE and FDE techniques [104]. All the DSP associated to the PDM CO-OFDM scheme described above can be generalized to an SDM CO-OFDM system.

The low-complexity equalization is however traded off against an overhead originating mainly from the cyclic prefix and the training symbols. The CP overhead is usually limited to 10 – 15% of

the useful OFDM symbol duration, which is able to completely address large accumulated dispersion values of tens of thousands of ps/nm without in-line DCFs (according to (2.12) and [134]). For SDM systems where modal dispersion can be larger, CP will introduce a significant overhead in terms of bandwidth and power. To cope with the large delays and limit the overhead, the CP can be set to a given size and if the experienced delay is larger than the CP, the extra delay can be dealt with using multi-tap time-domain filters to shorten the channel response before performing the OFDM demodulation [136]. This solution would still be less complex than SC-based time-domain equalization because the filters will only deal with a fraction of the total delay.

Even though today, no coherent OFDM optical products have been developed yet because SC systems were capable of providing 100 Gb/s and at the same time, OFDM was thought to be more sensitive to phase noise and to non-linear effects due to its high peak-to-average power ratio (PAPR), DSP algorithms were developed to protect the OFDM signals from phase noise and to compensate for non-linear effects. It has even been shown that at high data rates, OFDM systems have the same performance [124], or could even outperform [137] SC systems. From a broader perspective, with the growing demands for further high capacities, research works have shown that adoption of multi-carrier techniques, whether through OFDM or other super-channel formats cited in section 1.5, are highly recommended to reach high data rates. In wireless communications, OFDM demonstrates other advantages such as a spectrum partitioning tailored to frequency-selective channels which offers the capability of performing dynamic bit and power loading. Although this is not the case for most current optical channels (no severe frequency selectivity and no feedback to the transmitter), these advantages can be further investigated and exploited in the future to provide flexible and dynamically re-configurable optical networks.

2.4 MIMO coding techniques

Originally developed for wireless systems, MIMO signal processing is needed to retrieve the transmitted data by processing all the received data streams (from antennas in multi-antenna wireless channels, from polarization states in Pol-Mux systems or from modes and cores in SDM systems) at the same time. The simplest and most common equalizer is the previously presented ZF technique. However, this equalization method was found to be sub-optimal for wireless MIMO channels impaired by random fading effects, and more generally for all non-unitary MIMO channels. Hence, this also applies to optical MIMO schemes impaired by loss disparities and loss of orthogonality between its channels.

In this section, we shall see that more advanced DSP techniques can be implemented both at the transmitter and the receiver to better equalize the non-unitary effects in MIMO channels and/or reduce the performance penalties that they induce. In particular, we review MIMO coding

techniques developed for multi-antenna Rayleigh-fading wireless systems, and more specifically Space-Time coding techniques. These codes make the full use of all the dimensions, or degrees of freedom, of the MIMO channel to enhance the performance of the system. We recall the design criteria of these codes and present the major code families. Then, we move to the receiver side and list the various decoding techniques that can be applied to both, coded and uncoded MIMO schemes. These decoders can be classified into two categories: optimal and sub-optimal decoders. In other words, this section is a toolbox for MIMO processing that shows that only a sub-group of this toolbox was used in optical communications and that more techniques can be found to further enhance the performance.

2.4.1 Wireless MIMO channels

MIMO schemes emerged in wireless communications with the insertion of multiple antennas at the transmitter and the receiver as shown in Fig. 2.7, empowered by the well-known works of Foschini [135] and Telatar [138] who demonstrated that the channel capacity increases with the number of antennas in the system. This can be seen from the capacity expression of the square OFDM-MIMO channel given in (2.20). Let us consider a multi-antenna wireless system with n_t denoting the number of antennas at the transmitter and n_r the number of antennas at the receiver. The channel is supposed to be a Rayleigh fading frequency-flat channel that remains constant during T time slots (block-fading channel). The transmission can be modeled as:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N} \quad (2.24)$$

where $\mathbf{X}$ (resp. $\mathbf{Y}$) is the $n_t \times T$ (resp. $n_r \times T$) matrix of transmitted (resp. received) symbols during T time slots. $\mathbf{H}$ is a $n_r \times n_t$ matrix representing the channel where each term $h_{i,j}$ is the fading coefficient of the pathway between the i^{th} transmit antenna and the j^{th} receive antenna. These terms are independent and identically distributed (i.i.d.) complex Gaussian variables with zero mean and unit variance. Hence, at any time, there is a probability that the channel is in deep fading. $\mathbf{N}$ is a $n_r \times T$ matrix representing circularly symmetric white Gaussian noise $\sim CN(0, 2N_0\mathbf{I}_{n_r})$.

The MIMO system modeled in (2.24) offers simultaneously two major advantages: first, a data rate increase can be brought through spatial multiplexing over the n_t antennas. A popular transmitter architecture achieving this spatial multiplexing gain is the V-BLAST (Vertical-Bell Labs Space-Time Architecture). Second, a performance improvement can also be acquired by combating fading, using the spatial dimension (antennas). This advantage is commonly referred to as spatial diversity. In its largest sense, diversity refers to the transmission of the same information over different dimensions along with the reception of multiple independent observations of the replicas at the receiver. Diversity can operate over time, space, frequency or other dimensions.

Unfortunately, maximizing the multiplexing gain does not inherently lead to a maximization or even an access to the diversity gain. However, we know that block-fading multi-antenna MIMO schemes are bounded by a fundamental “Diversity-Multiplexing trade-off” defined by Zheng and Tse in [139]. To extract the spatial diversity of the MIMO channel and at the same time use its degrees of freedom for multiplexing, sophisticated coding across the transmit antennas is needed and Space-Time (ST) codes have been designed to achieve this trade-off for various MIMO systems.

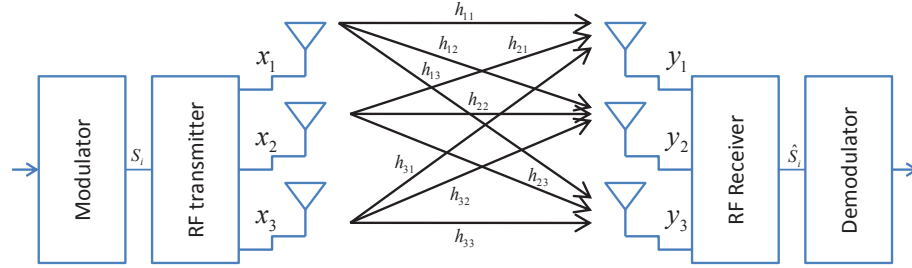


Figure 2.7: Multi-antenna wireless MIMO system.

2.4.2 Design criteria of Space-Time (ST) codes for Rayleigh fading channels

ST codes were first proposed by Tarokh et al. in [140] along with design criteria to build codes with an optimal performance on a MIMO Rayleigh block-fading channel, that were followed later by design criteria for fast-fading channels. To understand what is meant by optimum performance, we quickly review the design rules of ST codes considering the MIMO channel in (2.24). First, we suppose that the channel $\mathbf{H}$ is known or perfectly estimated at the receiver side - this is referred to as perfect channel state information (CSI) at the receiver - and let $\mathcal{C}$ be the set of all possibly transmitted matrices $\mathbf{X}$. Since the additive noise components are i.i.d., a maximum-likelihood (ML) decoder of (2.24) searches for the symbol matrix that is closest in Euclidean distance to the received matrix and outputs an estimate $\hat{\mathbf{X}}$ of $\mathbf{X}$ [11, Chap.15] according to:

$$\hat{\mathbf{X}} = \underset{\mathbf{X} \in \mathcal{C}}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_F^2 \quad (2.25)$$

where $\|\cdot\|_F$ is the Frobenius norm. The ML decoding criterion offers the lowest achievable BER if all symbol matrices $\mathbf{X}$ are equiprobable, which is generally the case.

The starting point for the construction of ST codes is the error probability of the transmission over the relevant MIMO channel defined as:

$$\begin{aligned} P_{\text{error}} &= \Pr\{\hat{\mathbf{X}} \neq \mathbf{X}\} \\ &= \sum_{\mathbf{X} \in \mathcal{C}} \Pr\{\mathbf{X}\} \Pr\{\hat{\mathbf{X}} \neq \mathbf{X} | \mathbf{X}\} \end{aligned} \quad (2.26)$$

which is the probability of sending a matrix $\mathbf{X}$ and detecting a different matrix $\hat{\mathbf{X}}$ averaged over all possibly transmitted symbol matrices. Using the union bound and assuming that the symbols are equiprobable, the error probability can be upper-bounded by [11, Chap.4,15]:

$$P_{\text{error}} \leq \frac{1}{\text{card}(\mathcal{C})} \sum_{\substack{\mathbf{X}, \hat{\mathbf{X}} \in \mathcal{C}, \\ \hat{\mathbf{X}} \neq \mathbf{X}}} \Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \quad (2.27)$$

where $\text{card}(\mathcal{C})$ is the cardinality of $\mathcal{C}$ and $\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}})$ is the pairwise error probability (PEP) that $\hat{\mathbf{X}}$ is selected when $\mathbf{X}$ was transmitted. When finding the PEP, we assume that only two symbol matrices exist in the set because the calculus of an exact P_{error} is intractable. The PEP for a given channel realization, or the conditional PEP is given by [11, Chap.15]:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}} | \mathbf{H}) = Q \left(\sqrt{\frac{\|\mathbf{H}(\hat{\mathbf{X}} - \mathbf{X})\|_F^2}{4N_0}} \right) \quad (2.28)$$

where Q is the Gaussian tail function given by:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp^{-t^2/2} dt \quad (2.29)$$

Using the Chernoff bound $Q(x) \leq \frac{1}{2} \exp^{-\frac{x^2}{2}} \quad \forall x$ and averaging the conditional PEP over the channel realizations, we obtain:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \mathbb{E}_H \left[\exp \left(-\frac{E_S \|\mathbf{H}\mathbf{E}\|_F^2}{8N_0} \right) \right] \quad (2.30)$$

where $\mathbb{E}_H[\cdot]$ is the averaging operation over all possible channel realizations, E_S is the symbol energy per transmit antenna and $\mathbf{E} = \frac{1}{\sqrt{E_S}} (\hat{\mathbf{X}} - \mathbf{X})$ is the normalized $n_T \times T$ error matrix. Now, we can average the PEP over the statistics of the Rayleigh fading channel which will yield the upper bound on the average PEP [11, Chap.15]:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \left(\prod_{n=1}^r \frac{1}{1 + \frac{E_S \lambda_n}{8N_0}} \right)^{n_r} \quad (2.31)$$

where r is the rank of the $n_t \times n_t$ matrix $\mathbf{A} = \mathbf{E}\mathbf{E}^\dagger$ and λ_n are the non-zero eigenvalues of $\mathbf{A}$. At high $\text{SNR} = \frac{E_S}{2N_0} \gg 1$, the PEP may be expressed as:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \left(\prod_{n=1}^r \lambda_n \right)^{-n_r} \left(\frac{\text{SNR}}{4} \right)^{-r \cdot n_r} \quad (2.32)$$

2.4.2.1 Diversity and coding gains

The expression of PEP for a Rayleigh slow fading MIMO channel in (2.32) shows three important parameters: n_r , r and $\{\lambda_n\}$. The number of antennas at the receiver n_r is fixed for a given system and shows a beneficial effect on the error probability that decays faster with an increased n_r . This is called the **receiver diversity**. On the other hand, r and $\{\lambda_n\}$ are determined by the properties of the transmitted matrices $\mathbf{X}$. With n_r , they define two criteria that improve the performance of the transmission: the rank criterion and the determinant criterion described by Tarokh *et al.* [140].

Rank criterion

The upper bound of the error probability in (2.32) decreases exponentially as a function of the SNR, and the product $r \cdot n_r$ determines the asymptotic slope of this decrease in a logarithmic plot. This product is called the **diversity order** of the MIMO system. The objective of the rank criterion is to achieve the maximum possible diversity of $n_t \cdot n_r$ obtained when the matrix $\mathbf{A}$ is full-rank ($r = n_T$) for any pair of different matrices in $\mathcal{C}$. When pure spatial multiplexing is performed, independent symbols are transmitted on each antenna and $\mathbf{X}$ is a $n_t \times 1$ vector. Hence, $r = 1$ and the maximum diversity cannot be reached. However if $\mathbf{X}$ is a $n_t \times T$ matrix with $T \geq n_t$, we may reach the full-rank condition. In fact $T = n_t$ is sufficient to obtain a full-rank $\mathbf{A}$ for all pairs and thus full diversity [140]. To fulfill this condition, $\mathbf{X}$ can be constructed by forming combinations $\{f_k\}$ of one or more modulated symbols s_k in each slot $x_{i,j}$ of $\mathbf{X}$ instead of sending independent data symbols. Accordingly, a possible codeword construction $\mathbf{X}_{ST}$ of a Space-Time code $\mathcal{C}$ of length T sent over n_t antennas can be:

$$\mathbf{X}_{ST} = \begin{bmatrix} f_1(s_1, s_2, \dots) & f_2(s_1, s_2, \dots) & \dots \\ f_3(s_1, s_2, \dots) & f_4(s_1, s_2, \dots) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (2.33)$$

$\mathbf{X}_{ST}$ has to satisfy a second criterion in order to achieve an optimum performance.

Determinant criterion

The determinant criterion follows from the first term on the right hand side of (2.32) that depends on the product of the non-zero eigenvalues of $\mathbf{A}$, hence the determinant of the largest non-singular square matrix in $\mathbf{A}$. The determinant criterion consists in maximizing the minimum of the determinants of the square matrices in $\mathbf{A}$ taken over all pairs of codewords. Indeed, the error probability of the system will be upper-bounded by the sum of the PEPs, however the PEP with the smallest determinant will be the dominant term in (2.27), and thus the minimum determinant should be maximized for an enhanced performance. The term involving the product of the non-zero eigenvalues in (2.32) is usually referred to as the coding gain.

Rate criterion

Both coding and diversity gains will be illustrated later in the chapter with specific ST codes. Before that, one additional design criterion of ST coding that does not explicitly appear in the error probability expression should be taken into account which is the **code rate**. A $n_t \times n_r$ MIMO channel can be seen as $\min(n_t, n_r)$ parallel AWGN channels (this can be verified by computing the capacity of the MIMO channel which will be a generalization of the one given in (2.20) for a square channel ($n_t = n_r$)). Hence, $\min(n_t, n_r)$ data symbols can be transmitted at each time slot. A $n_T \times T$ ST code is called a full-rate code if $\min(n_t, n_r) \cdot T$ symbols are sent during T time slots over the $n_t \times n_r$ channel.

In this section, we have presented three major design rules for ST codes in block Rayleigh fading channels. In the literature, other design criteria can be found based on, for instance, maximizing the mutual information between the transmitter and the receiver [141] yielding an independent criterion that does not systematically guarantee the maximum diversity. Another approach is suggested by Tse *et al.* in [142, Chap.9] where instead of averaging the PEP over all possible channel realizations, an outage formulation is adopted where the worst-case PEP over all channels that are not in outage is considered leading to a universal coding design criterion. Surprisingly, the resulting criterion is closely related to the determinant criterion suggesting that the latter is more fundamental than what it appears to be.

In the next section, we will focus on a popular category of ST codes which is ST Block Codes. However, it is noteworthy to mention that other coding techniques, already used for SISO channels, can be applied to MIMO channels such as the use of rotated QAM constellations [143] or conventional channel codes (FEC) with interleaving to achieve **signal or temporal diversity**. These schemes also include bit-interleaved coded systems [144] and trellis-coded modulations [145] for both SISO and MIMO as well as Space-Time trellis codes [140] for MIMO. However, we are mainly interested in ST codes that are tailored exclusively for MIMO schemes and that achieve the **spatial diversity** of the multi-antenna systems.

2.4.3 Space-Time Block Codes

Space-Time Block Codes (STBC) consist in $n_t \times T$ codeword matrices $\mathbf{X}_{ST}$ where T is the number of symbol time slots, also called channel uses (cu). The elements of $\mathbf{X}_{ST}$ are functions $f_{k=1 \dots n_t T}$ of $N_S \leq n_t T$ modulated symbols $s_1, s_2 \dots s_{N_S}$. A component $x_i(t_j)$ of the codeword matrix represents the new super-symbol transmitted over the i^{th} antenna and at the time slot t_j . An example of a 2×2 STBC code is given hereafter:

$$\mathbf{X}_{ST} = \begin{bmatrix} x_1(t_1) & x_1(t_2) \\ x_2(t_1) & x_2(t_2) \end{bmatrix} = \begin{bmatrix} f_1(s_1 \dots s_{N_S}) & f_2(s_1 \dots s_{N_S}) \\ f_3(s_1 \dots s_{N_S}) & f_4(s_1 \dots s_{N_S}) \end{bmatrix} \quad (2.34)$$

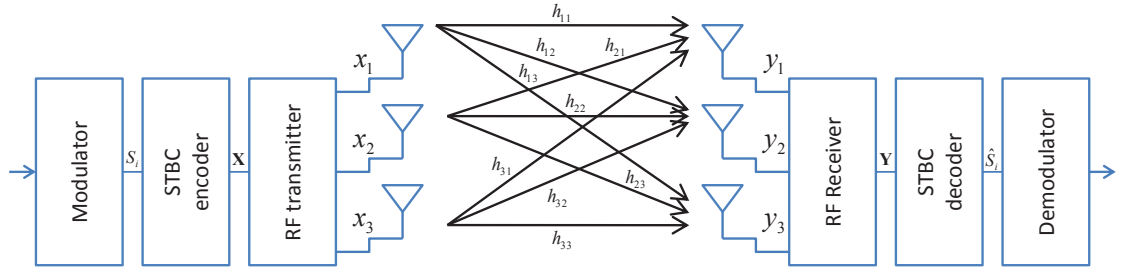


Figure 2.8: Space-Time block coded multi-antenna wireless MIMO system.

The rate of an STBC is the number of different modulated symbols sent per channel use:

$$r_{ST} = \frac{N_S}{T} (\text{symbols/cu}) \quad (2.35)$$

An STBC uses N_S symbols and generates a super-symbol matrix $\mathbf{X}_{ST}$ fed as a block to a set of n_t transmitters (antennas/polarizations/modes...) as can be seen in Fig. 2.8 for $n_t = n_r = 3$. When the functions f_k are linear combinations of the complex modulated symbols, the code is said to be complex linear and we can define a complex $n_t T \times N_S$ generator matrix $\mathbf{G}_C$ of the code by vectorizing, column-wise, the codeword matrix $\mathbf{X}_{ST}$ into an $n_t T \times 1$ vector such that:

$$\text{vec}_C(\mathbf{X}_{ST}) = \begin{bmatrix} x_1(t_1) \\ x_2(t_1) \\ \vdots \\ x_{n_t}(t_T) \end{bmatrix} = \mathbf{G}_C \mathbf{s}_{N_S \times 1} = \mathbf{G}_C \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{N_S} \end{bmatrix} \quad (2.36)$$

Another possible linear construction can involve both the modulated symbols and their conjugates. In this case, the STBC is characterized by a real $2n_t T \times 2N_S$ generator matrix $\mathbf{G}_R$, and is called a real linear code. The real vectorized model of the code is then given by:

$$\text{vec}_R(\mathbf{X}_{ST}) = \begin{bmatrix} \Re(x_1(t_1)) \\ \Re(x_2(t_1)) \\ \vdots \\ \Im(x_1(t_1)) \\ \Im(x_2(t_1)) \\ \vdots \end{bmatrix} = \mathbf{G}_R \mathbf{s}_R = \mathbf{G}_R \begin{bmatrix} \Re(s_1) \\ \Re(s_2) \\ \vdots \\ \Im(s_1) \\ \Im(s_2) \\ \vdots \end{bmatrix} \quad (2.37)$$

where $\text{vec}_R(\mathbf{X}) = [\Re(\text{vec}_C(\mathbf{X})) \ \Im(\text{vec}_C(\mathbf{X}))]^T$.

In order to guarantee that the linear system of equations between the transmitter and the receiver is not under-determined, the number of complex data symbols used to build the codeword is bounded by $N_S \leq \min\{n_t, n_r\} T$. Many STBC constructions exist in literature, some of them

being specific cases of larger constructions. In the following sections, we do not intend to cover all STBC families but we will rather focus on popular STBC families designed to answer three principal objectives:

- achieve the highest possible diversity in the MIMO system.
- achieve the highest possible rate, ideally being full-rate codes, i.e. sending $\min(n_t, n_r)$ data symbols per cu, thus causing no penalty in the spectral efficiency of the MIMO scheme.
- minimize the complexity of the STBC decoder at the receiver. Thus, complex and real linear codes will be considered as they can be optimally decoded with lattice decoders described in section 2.4.4.

2.4.3.1 Alamouti code

One of the first designed STBCs, Alamouti code [146] was tailored for 2×1 MIMO systems on which it achieves a full rate of 1 symbol/cu and full diversity. The codeword matrix of the Alamouti code is the following:

$$\mathbf{X}_A = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \quad (2.38)$$

where s_1 and s_2 are two modulated symbols. The codeword matrix has an orthogonal structure satisfying $\mathbf{X}_A \mathbf{X}_A^\dagger = (|s_1|^2 + |s_2|^2) \mathbf{I}_2$ which makes its decoding straightforward [146]. A 2×1 Alamouti-coded scheme is given by:

$$\begin{bmatrix} y_1(t_1) & y_1(t_2) \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} + \begin{bmatrix} n_1 & n_2 \end{bmatrix} \quad (2.39)$$

and can be rewritten as:

$$\underbrace{\begin{bmatrix} y_1(t_1) \\ y_1(t_2)^* \end{bmatrix}}_{\tilde{\mathbf{y}}} = \underbrace{\begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}}_{\tilde{\mathbf{H}}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \underbrace{\begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}}_{\tilde{\mathbf{n}}} \quad (2.40)$$

In this equivalent model $\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{n}}$, $\tilde{\mathbf{H}}$ has the same structure of the Alamouti codewords and is hence orthogonal (or unitary). Applying a simple channel inversion using $\tilde{\mathbf{H}}^\dagger$ followed by a threshold decision (ZF decoding) when $\mathbf{H}$ is unitary is the same as applying the ML criterion in (2.25). Accordingly, we get:

$$\tilde{\mathbf{H}}^\dagger \tilde{\mathbf{y}} = \begin{bmatrix} \|h_1\|^2 + \|h_2\|^2 & 0 \\ 0 & \|h_1\|^2 + \|h_2\|^2 \end{bmatrix} \mathbf{s} + \tilde{\mathbf{H}}^\dagger \tilde{\mathbf{n}} \quad (2.41)$$

where the components of the noise vector are not amplified but remain Gaussian with the same variance as the original channel noise due to the orthogonal structure of the equivalent channel.

Hence, the ML criterion is satisfied through a simple linear processing which explains the popularity of this ST code (known as matrix A in wireless standards). The Alamouti code can also be applied on a 2×2 MIMO channel. However, it introduces a significant redundancy as it becomes a half-rate code achieving a rate of only 1 symbol/cu whereas the maximum rate in this case is 2 symbols/cu.

Orthogonal codes

In [147], Tarokh *et al.* proved that orthogonal codes exist for larger MIMO schemes with a simple linear-decoding advantage at the receiver while achieving an optimal ML performance thanks to the following property:

$$\mathbf{X}\mathbf{X}^\dagger = \left(\sum_{i=1}^{n_t} |s_i|^2 \right) \mathbf{I} \quad (2.42)$$

where s_i are the data symbols. However, the rate of these orthogonal codes is less than 1 symbol/cu for $n_t > 2$ and decreases with an increasing n_t . For K transmitted data symbols, the maximum rate of an orthogonal code is $r = \frac{K}{2^{K-1}}$ [147]. A codeword example of an orthogonal code for 3 transmit antennas with a rate $r = 3/4$ is given below:

$$\mathbf{X} = \begin{bmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & -s_3 \\ -s_3^* & 0 & s_1^* & s_2 \end{bmatrix} \quad (2.43)$$

2.4.3.2 Threaded Algebraic Space Time (TAST) codes

The limited rate of orthogonal codes motivated the search for new spectral-efficient linear code constructions. Among the first systematic constructions of full-rate and full-diversity STBCs, the Threaded Algebraic Space-Time (TAST) code construction proposed by El Gamal and Damen in [148] is the most popular for its generality, simplicity and performance. TAST codes make use of component codes designed for SISO channels. Each component code is assigned to a thread in the codeword and each thread is spread on different antennas at different time slots. The threaded structure can be seen in Fig. 2.9 for a 3×3 and a 6×6 TAST code. The corresponding codeword matrices will be explicitly given in chapter 4.

TAST codes are designed in two steps. First, each isolated thread is guaranteed to achieve full-diversity by rotating q -QAM symbols using appropriate $n_t \times n_t$ rotation matrices [143]. Second, to make sure that the stacked threads form a full-diversity code, they are scaled by a set of algebraic numbers or transcendental numbers [148] that provides an algebraic separation of the threads. However, the TAST codewords have a determinant, and hence a coding gain, that vanishes with higher-order modulation symbols which prevents them from achieving the optimal diversity-multiplexing trade-off.

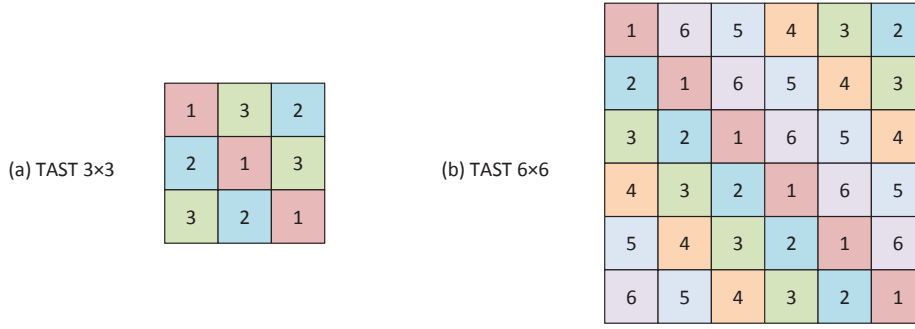


Figure 2.9: Threaded Space-Time architecture for $n_t = n_r = \{3, 6\}$.

2.4.3.3 Perfect codes

Perfect codes [149] have the same threaded structure as TAST codes. They also achieve a full-rate and a full-diversity of $n_t \cdot n_r$ as well as the optimal diversity-multiplexing trade-off over Rayleigh fading channels. A judicious choice of the rotation matrices within each thread along with a thread separation based on cyclic division algebras ensure a coding gain that remains constant with an increased constellation size (a property called: **non-vanishing determinant**). All these attractive properties along with a uniform average energy transmitted per antenna warrant them the “perfect codes” label. These codes are also popular for their generality, they were originally proposed for 2,3,4 and 6 transmit antennas [149] and later extended to any number of antennas [150]. A famous perfect code is the 2×2 code, also known as the **Golden code** invented by Belfiore *et al.* [151] and called so for its use of the Golden number $\frac{1+\sqrt{5}}{2}$.

Golden code

The Golden code [151] has the best performance on 2×2 MIMO Rayleigh fading channels. The codeword matrix of the Golden code is:

$$\mathbf{X}_G = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha(s_1 + \theta s_2) & \alpha(s_3 + \theta s_4) \\ i\bar{\alpha}(s_3 + \bar{\theta} s_4) & \bar{\alpha}(s_1 + \bar{\theta} s_2) \end{bmatrix} \quad (2.44)$$

where $\theta = \frac{1+\sqrt{5}}{2}$, $\bar{\theta} = \frac{1-\sqrt{5}}{2}$, $\alpha = 1 + i - i\theta$, $\bar{\alpha} = 1 + i - i\bar{\theta}$ and s_1, s_2, s_3, s_4 are four modulated symbols. The Golden code achieves a full rate of 2 symbols/cu because 4 different symbols are transmitted from 2 antennas during 2 symbol times. Hence, no redundancy is introduced by the code. Moreover, the determinant of the codeword matrix corresponding to a coding gain on a Rayleigh fading channel is equal to $\frac{1}{5}$ which is the highest possible value for a 2×2 STBC [151].

The Golden code is complex linear with a unitary generator matrix given by:

$$\mathbf{G}_{\mathbb{C},\mathcal{G}} = \begin{bmatrix} \alpha & \alpha\theta & 0 & 0 \\ 0 & 0 & i\bar{\alpha} & i\bar{\alpha}\bar{\theta} \\ 0 & 0 & \alpha & \alpha\theta \\ \bar{\alpha} & \bar{\alpha}\bar{\theta} & 0 & 0 \end{bmatrix} \quad (2.45)$$

The constellation of the Golden codewords built using 4-QAM symbols is given in Fig. 2.10.

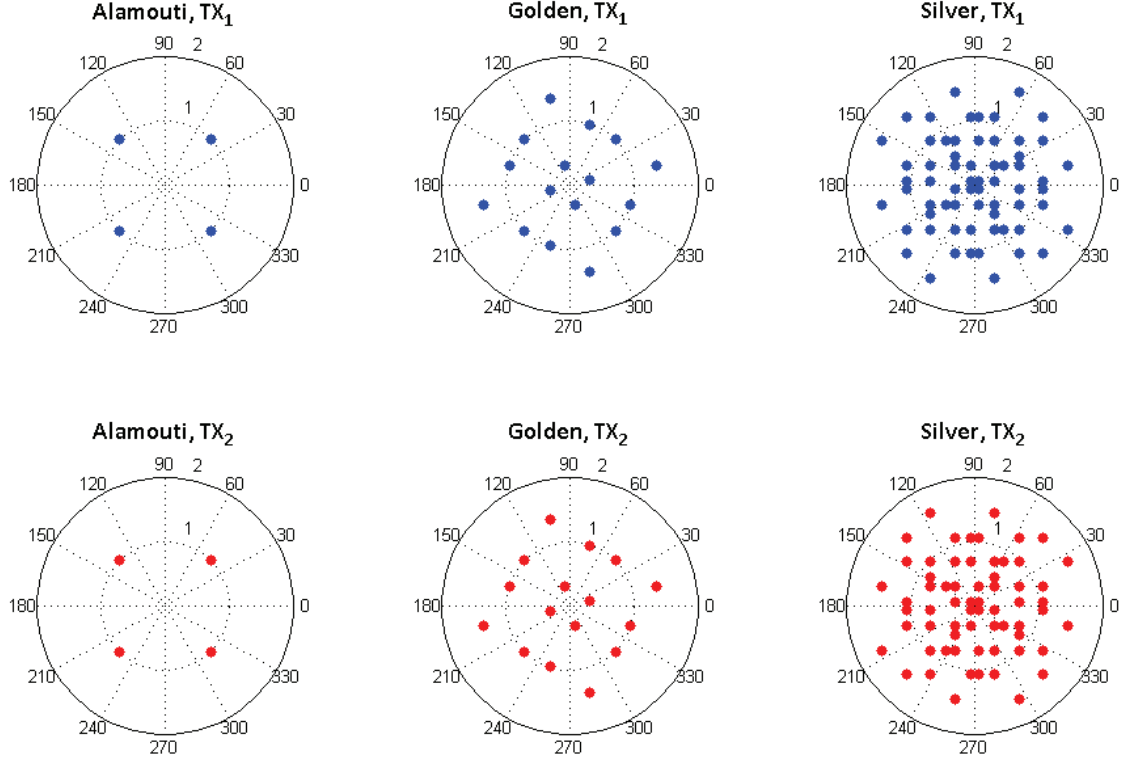


Figure 2.10: Constellations of STBCs using 4-QAM symbols: Alamouti, Golden and Silver codes.

2.4.3.4 Silver code

While the Golden code has the best performance over a 2×2 Rayleigh fading MIMO channel, other code constructions were still investigated in which performance can be traded-off for a lower-complexity decoding. The Silver code developed by Tirkkonen *et al.* is a full-rate and full-diversity code that has a slightly weaker performance than the Golden code, hence its name: Silver code. The codeword matrix is based on a layered structure of two Alamouti codes [152] which makes it a fast-decodable code [153]:

$$\mathbf{X}_S = \frac{1}{\sqrt{2}} \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} z_1 & -z_2^* \\ z_2 & z_1^* \end{bmatrix} \quad (2.46)$$

with $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \frac{1}{\sqrt{7}} \begin{bmatrix} 1+i & -1+2i \\ 1+2i & 1-i \end{bmatrix} \begin{bmatrix} s_3 \\ s_4 \end{bmatrix}$

where s_1, s_2, s_3, s_4 are four modulated symbols. The determinant of this code is equal to $\frac{1}{7}$, slightly inferior to the one achieved by the Golden code. The Silver code is real linear with the following unitary generator matrix:

$$\mathbf{G}_{\mathbb{R},S} = \frac{1}{\sqrt{(2)}} \begin{bmatrix} 1 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{7}} & 0 & 0 & -\frac{1}{\sqrt{7}} & -\frac{2}{\sqrt{7}} \\ 0 & 1 & -\frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{7}} & 0 & 0 & \frac{2}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \\ 0 & -1 & -\frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{7}} & 0 & 0 & \frac{2}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \\ 1 & 0 & -\frac{1}{\sqrt{7}} & \frac{1}{\sqrt{7}} & 0 & 0 & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{7}} \\ 0 & 0 & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{7}} & 1 & 0 & \frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \\ 0 & 0 & -\frac{2}{\sqrt{7}} & \frac{1}{\sqrt{7}} & 0 & 1 & -\frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \\ 0 & 0 & \frac{2}{\sqrt{7}} & -\frac{1}{\sqrt{7}} & 0 & 1 & \frac{1}{\sqrt{7}} & \frac{1}{\sqrt{7}} \\ 0 & 0 & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{7}} & -1 & 0 & \frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \end{bmatrix} \quad (2.47)$$

The constellation of the Silver codewords built using 4-QAM symbols is also given in Fig. 2.10. The BER performance of ML-decoded uncoded, Golden- and Silver-coded transmissions over a 2×2 Rayleigh fading channel are plotted in Fig. 2.11. While the uncoded spatial-multiplexed scheme (blue curve) has a diversity order of $r \cdot n_r = 2$ corresponding to the number of antennas at the receiver, the Silver- and Golden- coded schemes have a diversity order of $r \cdot n_r = 4$ which is the maximum diversity of the 2×2 MIMO channel. The diversity order can be graphically discerned from the slopes of the average BER curves that decay faster when the codes are used. Apart from the diversity gain, the codes horizontally shift the BER curves to the left which translates in a coding gain. As predicted from the determinant criterion, the Golden code shows a better coding gain than the Silver code.

2.4.4 Decoders for MIMO schemes

We have presented the ML decoding criterion in (2.25) and derived the corresponding PEP of a Rayleigh block-fading frequency-flat MIMO channel. ML decoding is the strategy yielding the best BER performance over the linear MIMO channel in (2.24) assuming that the transmitted codewords are equiprobable. In section 2.4.2, we highlighted a spatial diversity gain of n_r for simple spatial multiplexing and $n_t \cdot n_r$ for a full-rank ST code when ML decoding is performed. The performance of the ML-decoded schemes over a Rayleigh fading MIMO channel in Fig. 2.11 showed that this decoder achieves the full diversity of each system.

Had we considered a ZF decoding (channel inversion followed by a decision threshold), a diversity order of only $n_r - n_t + 1$, when $n_r \geq n_t$, would have been achieved [11, Chap.15]. This is illustrated in Fig. 2.11 for the 2×2 channel where ZF decoding of an uncoded scheme yields a BER curve decreasing inversely as the SNR raised to the power 1 compared to the power 2 achieved by an ML decoder. The diversity order of the ZF equalizer can be explained by the fact that the decoder has n_r degrees of freedom provided by the n_r antennas at the receiver and through

its channel inversion, it is using $n_t - 1$ degrees of freedom to cancel the interference of the $n_t - 1$ signals from each detected data symbol, resulting in a diversity of $n_r - (n_t - 1)$. We also show in Fig. 2.11 that a ZF decoding of an STBC, in this case the Golden code, brings neither a diversity nor a coding gain. The procurement of these gains is conditioned by the use of an ML decoder.

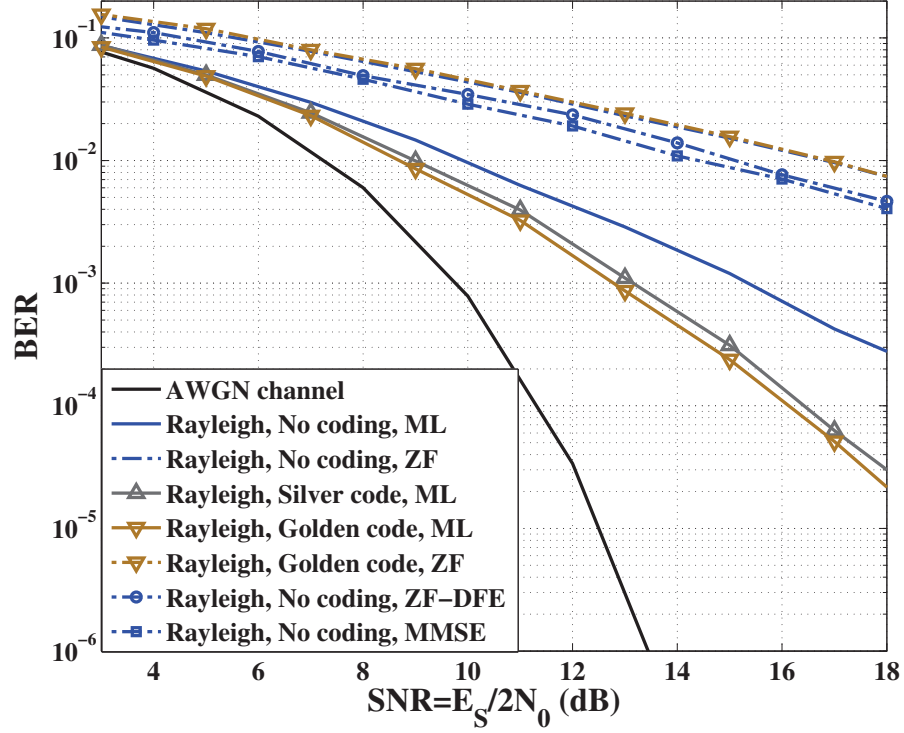


Figure 2.11: Golden and Silver code performance on 2×2 Rayleigh fading channel with ML decoding. The performance of some sub-optimal decoders of the uncoded and the Golden-coded schemes is also shown. The AWGN-channel performance is given as a reference.

After this overview on ST coding, we dedicate the following section to the review of some efficient techniques to implement the ML criterion called “optimal decoders”. We also present two of the most popular lower-complexity sub-optimal decoders for both uncoded and STBC-coded MIMO channels.

2.4.4.1 Optimal decoders

From (2.25), the straightforward way to implement the ML criterion is an exhaustive search in the codeword set $\mathcal{C}$. For each $\mathbf{X}$, the Euclidean norm $\|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2$ has to be computed. Given the linear property of the considered STBCs (real or complex linear), the ML criterion can be further developed to explicitly show the original data symbols when an STBC scheme is used. To this end, we use a vectorized form of (2.24) containing the generator matrix of the ST code and define

the equivalent channel $\mathbf{H}_{eq}$ [154] such that:

$$\begin{aligned} \mathbf{y}'_{n_r T \times 1} = \text{vec}_{\mathbb{C}}(\mathbf{Y}) &= \begin{bmatrix} \mathbf{H} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H} \end{bmatrix} \text{vec}_{\mathbb{C}}(\mathbf{X}) + \text{vec}_{\mathbb{C}}(\mathbf{N}) \\ &= \underbrace{\mathbf{H}'_{n_r T \times n_t T} \mathbf{M}_G}_{\mathbf{H}_{eq}} \mathbf{s}_{n_t T \times 1} + \mathbf{n}'_{n_r T \times 1} = \mathbf{H}_{eq} \mathbf{s} + \mathbf{n}' \end{aligned} \quad (2.48)$$

where $\mathbf{M}_G$ is the generator matrix of the coding scheme and $\mathbf{H}_{eq}$ the new $n_r T \times n_t T$ equivalent channel. In the case of simple spatial multiplexing without coding, $T = 1$ and $\mathbf{M}_G$ is replaced by the identity matrix. $\mathbf{s} = [s_1, \dots, s_i, \dots, s_{n_t T}]$ are the transmitted data symbols. Assuming that $\mathbf{H}$ is a full-rank matrix and $\mathbf{M}_G$ is unitary for the investigated full-rank ST codes, the ML decoding rule can be reinterpreted as:

$$\hat{\mathbf{s}}_{ML} = \underset{\mathbf{s}_{n_t T \times 1} \in \mathcal{C}'}{\text{argmin}} \|\mathbf{y}' - \mathbf{H}_{eq} \mathbf{s}\|^2 \quad (2.49)$$

where $\mathcal{C}'$ is the set of all possible transmitted q-QAM symbols. An exhaustive search in $\mathcal{C}'$ requires q^{n_t} norm computations to decode a spatially multiplexed scheme ($T = 1$) and $q^{n_t^2}$ norm computations to decode a full-rate square STBC ($T = n_t$) that uses q -QAM modulated symbols. The complexity of the exhaustive search increases with the constellation size and grows exponentially with the number of antennas (or spatial modes). Hence, lower-complexity decoders implementing the ML criterion are needed.

Lattice decoders

After applying a complex-to-real transformation of the channel in (2.48) and assuming that the used modulation format is q-QAM ($s_i \in \mathbb{Z}^2$), we can rewrite (2.48) as:

$$\begin{aligned} T(\mathbf{y}') &= \begin{bmatrix} \Re(\mathbf{H}_{eq}) & -\Im(\mathbf{H}_{eq}) \\ \Im(\mathbf{H}_{eq}) & \Re(\mathbf{H}_{eq}) \end{bmatrix} T(\mathbf{s}) + T(\mathbf{n}') \\ \mathbf{y}'_{\mathbb{R}} &= \mathbf{H}_{eq, \mathbb{R}} \mathbf{s}_{\mathbb{Z}} + \mathbf{n}'_{\mathbb{R}} \end{aligned} \quad (2.50)$$

where $T(\cdot) = [\Re(\cdot) \ \Im(\cdot)]^T$. Given the full-rank assumption of the equivalent channel $\mathbf{H}_{eq}$, $\mathbf{H}_{eq, \mathbb{R}}$ forms a basis of $\mathbb{R}^n$ with $n = 2n_t T$. Thus, it can be associated with a lattice $\Lambda_{\mathbf{H}_{eq}}$ obtained by a linear transformation of $\mathbb{Z}^n$. A lattice in $\mathbb{R}^n$ is a discrete subgroup of $\mathbb{R}^n$ that spans the real vector space in $\mathbb{R}^n$ and is given by the set:

$$\Lambda = \left\{ \sum_{i=1}^n a_i \mathbf{v}_i \mid a_i \in \mathbb{Z} \right\} \quad (2.51)$$

where $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ forms a basis of $\mathbb{R}^n$. Hence, the ML decoding rule can be rewritten as:

$$\hat{\mathbf{s}}_{\mathbb{Z},ML} = \underset{\mathbf{s}_{\mathbb{Z},n \times 1} \in \mathbb{Z}^n}{\operatorname{argmin}} \|\mathbf{y}'_{\mathbb{R}} - \mathbf{H}_{eq,\mathbb{R}} \mathbf{s}_{\mathbb{Z}}\|^2 \quad (2.52)$$

and the search for the ML solution brings us to solving a closest-vector problem, i.e. finding the lattice-point that is closest to $\mathbf{y}'_{\mathbb{R}}$, which can be implemented using improved search strategies.

Reduced-search lattice decoders such as the sphere decoder were suggested in [155, 156] for both uncoded and ST-coded systems. Due to the propagation through a noisy channel, the received vector $\mathbf{y}'_{\mathbb{R}}$ is no longer a lattice point and the sphere decoder will search for the closest lattice point in a finite hyper-spherical region centered on $\mathbf{y}'_{\mathbb{R}}$. Given that the number of visited point in the finite hyper-sphere is limited, the decoding complexity is reduced. If, for a certain initial radius, the sphere was found to be empty, the radius is increased and the closest lattice point found in the new sphere, resulting hence in an optimal ML performance. The choice of the radius C_{SD} of the sphere is crucial for the complexity reduction of the search algorithm. If the radius is judiciously chosen in function of the noise power and the properties of the channel according to [156]:

$$C_{SD}^2 = \min [\min \{\lambda_{i=1 \dots 2n}\}, 2nN_0] \quad (2.53)$$

where λ_i are the singular values of $\mathbf{H}_{eq,\mathbb{R}} \mathbf{H}_{eq,\mathbb{R}}^T$, the complexity of the search algorithm becomes independent of the constellation size q and is approximated by $\mathcal{O}((n_t T)^6)$ [156]. This gives $\mathcal{O}(n_t^6)$ operations for uncoded spatial multiplexing and $\mathcal{O}(n_t^{12})$ for a full-rate square ST code. In appendix D, we describe the search algorithm of the sphere decoder presented in [155, 156].

A lattice decoder with a different search strategy inside spherical regions is the Schnorr-Euchener decoder [157]. Furthermore, the stack decoder [158] is also a famous reduced-search decoder based on a more general branch-and-bound tree search technique and a sequential decoding process. It progressively computes the Euclidean norm in the ML criterion and orders the most promising symbols minimizing this metric in a list called a “stack”. The stack decoder can be associated to a spherical bound limiting the search region which yields even more computational savings than the original stack or sphere decoders [159].

2.4.4.2 Sub-optimal decoders

Until now, we have seen one sub-optimal linear MIMO decoder which is the ZF equalizer in (2.18) that consists in an inverse channel detector. We have noted that this decoding technique does not exploit the diversity of the MIMO channel and results in an error probability with a lower slope. A major problem of the ZF equalizer is noise amplification. Indeed, by applying the inverse of the channel on the received signal, the deep fading channel coefficients (low-values of $h_{i,j}$) amplify the additive noise and lead to more detection errors. This is avoided when the MIMO channel

is unitary, i.e. $\mathbf{H}\mathbf{H}^\dagger = \mathbf{I}$ and the properties of the additive noise are not affected by the channel inversion. Hence, the noise remains white and Gaussian. Only in this case, a ZF decoding becomes optimal, i.e. achieves the ML criterion in (2.25).

Other sub-optimal MIMO decoders were suggested to palliate the drawbacks of ZF equalization while maintaining a low complexity. We will present hereafter two sub-optimal decoders: the minimum-mean square error (MMSE) detector and a ZF with decision feedback equalizer (ZF-DFE).

The MMSE equalizer linearly combines the received signals with a filter $\mathbf{W}_{MMSE}$ chosen to minimize the following mean square error applied over the real equivalent system in (2.52):

$$\mathbf{W}_{MMSE} = \underset{\mathbf{W}}{\operatorname{argmin}} \mathbb{E}(\|\hat{\mathbf{s}}_{\mathbb{R},MMSE} - \mathbf{s}_{\mathbb{Z}}\|^2) = \underset{\mathbf{W} \in \mathbb{R}^{2n_r T \times 2n_t T}}{\operatorname{argmin}} \mathbb{E}(\|\mathbf{W}\mathbf{y}'_{\mathbb{R}} - \mathbf{s}_{\mathbb{Z}}\|^2) \quad (2.54)$$

The linear MMSE filter achieving this minimization is found to be [11, Chap.15]:

$$\mathbf{W}_{MMSE} = \left(\mathbf{H}_{eq,\mathbb{R}}^T \mathbf{H}_{eq,\mathbb{R}} + \frac{1}{\text{SNR}} \mathbf{I} \right)^{-1} \mathbf{H}_{eq,\mathbb{R}}^T \quad (2.55)$$

$\mathbf{W}_{MMSE}$ is hence applied over the received vector $\mathbf{y}'_{\mathbb{R}}$ to get $\hat{\mathbf{s}}_{\mathbb{R},MMSE}$ on which a threshold decision is performed to estimate the data symbols. The MMSE detector outperforms ZF decoding as can be seen in Fig. 2.11. Although it achieves the same diversity order as ZF equalization, the inverse SNR term in its linear filter limits noise amplification at low SNR values and outperforms ZF.

Assuming $\mathbf{H}_{eq}$ is a full-rank matrix with $n_r \geq n_t$, The ZF-DFE is a non-linear equalization performed as follows:

- Start with a QR decomposition of $\mathbf{H}_{eq,\mathbb{R}} = \mathbf{Q}\mathbf{R}$ that rewrites the channel matrix as the product of $\mathbf{Q}$, a unitary matrix and $\mathbf{R}$, an upper triangular matrix (refer to appendix C for more details on the QR decomposition).
- Compute the equivalent system:

$$\tilde{\mathbf{y}}_{\mathbb{R}} = \mathbf{Q}^T \mathbf{y}'_{\mathbb{R}} = \mathbf{R}\mathbf{s} + \mathbf{Q}^T \mathbf{n}' \quad (2.56)$$

where the new noise $\mathbf{Q}^T \mathbf{n}'$ remains white Gaussian given that $\mathbf{Q}$ is a unitary matrix.

- Estimate each symbol in the vector $\hat{\mathbf{s}}_{ZF-DFE}$ by solving the linear system $\tilde{\mathbf{y}} = \mathbf{R}\mathbf{s}$. The matrix $\mathbf{R}$ being upper triangular, the system can be solved in an iterative fashion starting from the last symbol and performing a threshold decision on each estimated symbol before

feeding it to the previous equation. This translates into:

$$\begin{aligned} s_{i,i:n \rightarrow 1} &= \frac{y_i - \sum_{j=0}^{i-1} r_{i,n-j} \hat{s}_{n-j}}{r_{i,i}} \\ \hat{s}_{i,i:n \rightarrow 1} &= \lfloor s_i \rfloor \end{aligned} \quad (2.57)$$

where $\lfloor \cdot \rfloor$ is a threshold decision operation.

The non-linearity comes from the symbol-wise threshold decision operation. The ZF-DFE demonstrates a performance gain over the classic ZF decoder through its successive interference cancellation while retrieving the data symbols [160]. However, ZF-DFE is prone to error propagation in its iterative decoding structure. Enhanced versions of ZF-DFE include reordering the channel matrix in order to decode the most reliable received symbols, i.e. having the highest received SNR, first.

Summary

In this chapter, the main digital signal processing tools that improved both the reach and the performance of the installed high-capacity optical fiber transmission systems were reviewed. Then, we focused on advanced DSP and coding techniques for the emerging MIMO schemes in optical fiber links. In particular, we identified OFDM as a promising modulation format for future low-complexity, flexible and robust systems. Later, we focused on Space-Time codes, initially designed for multi-antenna wireless systems and more recently used in [8] for PDL-impaired polarization-multiplexed optical systems. Design rules of ST codes, families of ST block codes as well as several decoding techniques for both ST-coded and uncoded MIMO channels were recollected. In the following chapters, the benefits of implementing ST codes in optical MIMO transmission systems will be further analyzed. When needed, we will refer to the codeword matrices of the investigated codes and the decoding algorithms presented in this chapter.

Chapter 3

Polarization-Time Coding for PDL Mitigation in Polarization Multiplexed Systems

Our first application of Space-Time coding techniques in optical MIMO channels is PDL mitigation in polarization-multiplexed or PDM systems. The polarization dimension has certainly attracted a lot of attention in coherent optical communications, its main perceived benefit being the transmission of independent data streams on a set of two orthogonal polarizations, thus doubling the capacity compared to a single-polarization scheme. However, optimized usages of this dimension were later investigated in order to improve the reach of the system. In [39], Agrell and Karlsson suggested using the polarization and the quadrature dimensions I_X , Q_X , I_Y , Q_Y to implement four-dimensional (4D) modulation formats with an increased OSNR sensitivity (in other words, a larger minimal distance d_{min} between the constellation points) compared to QPSK over an AWGN channel. An example of 4D modulation is polarization-switched QPSK (PS-QPSK) which increases the minimal distance at the cost of a reduced spectral efficiency because this format has only 8 constellation points compared to 16 constellation points for PDM-QPSK [39]. In 2013, a capacity-distance product record-breaking transmission using 4D-16QAM coded modulations instead of multiplexing two 16-QAM symbols over a PDM system was demonstrated [113], trading spectral efficiency with an enhanced sensitivity. In our work, we propose using the polarization dimension to serve two purposes: multiplexing data and protecting it from any gain disparities or loss of orthogonality between the data-carrying polarizations, both induced by polarization dependent loss (PDL) as seen in chapter 1. ST coding is a technique that can fulfill these two objectives.

The benefits of applying ST Block Codes (STBCs) for PDL mitigation in PDM-OFDM systems were shown through preliminary numerical simulations in a previous thesis work carried in

our labs by Mumtaz *et al.* [8]. While OFDM manages the dispersion in the channel, STBCs deals with PDL and reduces the induced penalties. Later, STBCs were also applied in single-carrier PDM systems showing similar results but necessitating a more complex decoding with larger time-domain filters [161]. In [162], a simple space-time trellis coding scheme is suggested for enhancing the tolerance of single-carrier PDM systems to PDL, inspired from an optical birefringent element that spreads the data-carrying polarizations over different states, thus averaging the effects of PDL. However, its performance is inferior to that of the investigated STBCs in previous works.

In this chapter, we recall the findings of [8] in section 3.1 and build upon them, extending the simulations to a complete PDM-OFDM system and a more elaborate channel model in order to test the resilience of ST codes to the particularities of the optical channel such as the statistics of PDL and the presence of non-linear effects in section 3.2. Later on in section 3.3, we analyze the observed performance by looking into the pairwise error probability (PEP) of an optical channel with PDL. We derive an upper bound of the PEP that can be used to predict how optimal a ST coding scheme is in combating MDL. The upper bound also yields a design criterion of optimal ST codes for PDL mitigation. Finally in section 3.4, we present an experimental demonstration of the coding gains brought by ST codes in a 1000 km PDL-impaired PDM transmission carried out in our laboratories.

3.1 Polarization-Time coding: preliminary results

A clear description of the MIMO modeling of coherent PDM transmission systems is given in [163] where authors emphasize that many MIMO techniques, including ST codes, originally developed for wireless channels can be applied in optical MIMO channels to mitigate linear and non-linear effects. In [164], the term “Polarization-Time coding” (or PT coding) was used to describe an Alamouti code-based coherent optical transmission system proposed to mitigate polarization scattering induced by non-linear XPM in WDM systems. Later, in [165], Alamouti code was also proposed in order to mitigate PMD in coherent optical OFDM systems.

Two main observations can be made from these works: first, the Alamouti-code based solutions systematically reduce in half the multiplexing gain that can be achieved on a 2×2 optical MIMO system leading to a sub-optimal use of the available degrees of freedom. Second, concerning PMD mitigation, we have seen in chapter 2 various frequency-domain and time-domain equalization techniques able to efficiently remove dispersive effects including PMD that is by nature a unitary effect that modifies the state of polarization of the optical signal without altering its total energy. Indeed, a fiber span represented by a concatenation of PMD elements as described in chapter 1 results in a unitary transform matrix $\mathbf{F}_k$ per OFDM subcarrier (i.e. $\mathbf{F}_k \mathbf{F}_k^\dagger = \mathbf{I}$), and hence

induces no fading. PMD might be a limiting effect especially for direct-detection systems, however it can be compensated for at the receiver in coherent-detection systems and does not reduce the capacity of the optical 2×2 channel. Given that the singular values of $\mathbf{F}_k$ are equal to 1, the capacity computed as in (2.20) corresponds to the one of two parallel AWGN channels. Hence, PT coding would bring no gains in a PMD-impaired coherent transmission system.

On the other hand, the transfer matrix of PDL in (1.21) reveals a non-unitary fading-like effect introduced by distributed low-PDL optical components. In long-haul links, PDL can accumulate leading to a capacity reduction and a performance degradation. ST coding might be an interesting technique to mitigate PDL-induced fading effects. However, the nature of this fading differs from the Rayleigh fading where the matrix coefficients are i.i.d. zero-mean Gaussian variables. Indeed, the periodic optical amplification preserving the total optical power in long-distance links prevents the fading of both polarizations at the same time.

The efficiency of ST coding in mitigating PDL was actually observed in previous works [8, 166–168]. In [8], PDL-mitigation capability of PT codes was investigated considering a single OFDM subcarrier and a simple PDL model using the matrix $\mathbf{H}_{PDL}$ in (1.21). The performance of three codes: Alamouti, Silver and Golden codes was simulated revealing that the Silver code outperformed the Golden code which is different from the performance over a Rayleigh fading channel where the Golden code is optimal. The Silver code removed nearly all the PDL-induced power penalty for PDL values up to 6 dB. Moreover, in [168], the Alamouti code exhibited the same performance independently of the PDL amount Γ_{dB} . In [166], an outage analysis showed that the Silver code was optimal for the low to moderate PDL values (up to 5 dB), whereas, in the case of large PDL, Alamouti code showed the best performance.

3.2 Numerical study of a Polarization-Time coded optical OFDM transmission

We start by extending the simulation work in [8] to a complete OFDM transmission and a long-haul optical link in which dispersive effects (CD and PMD) are taken into account. Having a low-complexity advantage over other equalization formats as denoted in sections 2.2 and 2.3, OFDM will be considered in the rest of the thesis. Applications of ST coding in the context of a single-carrier optical transmission were also investigated in literature [161, 162]. They require, however, more complex channel equalizers.

The scheme of a PT-coded PDM OFDM system is shown in Fig. 3.1 where the OFDM transmitter contains the same blocks as in Fig. 2.4. An OFDM signal is sent on each polarization. Independently modulated data symbols are assigned to the different orthogonal subcarriers of each OFDM symbol when no PT code is considered, whereas the data symbols are linearly combined

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

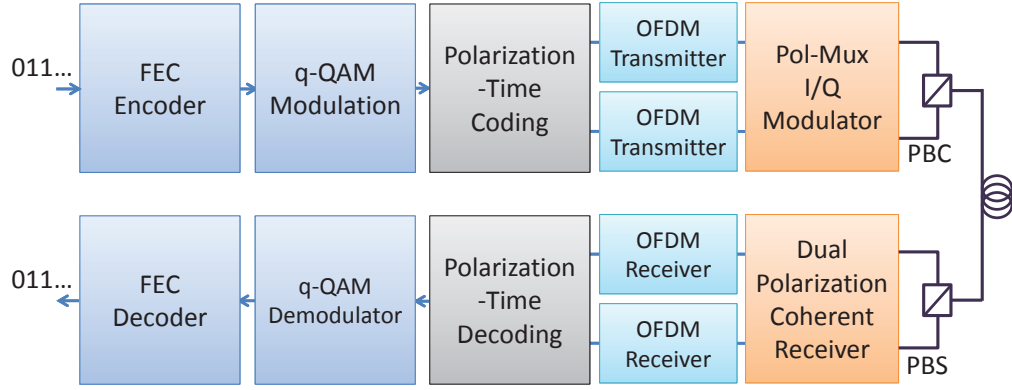


Figure 3.1: General scheme of a PDM-OFDM transmission system with PT coding.

Table 3.1: Simulation parameters related to the optical channel on the left and to the OFDM modulation format on the right

Total length L_{tot}	20×100 km	Total bit-rate	40 Gb/s
Transmission loss	0.2 dB/km	OFDM subcarriers	256
CD coefficient D	17 ps/nm/km	OFDM pilot tones N_p	5
PMD coefficient PMD	$0.05 - 1$ ps/ $\sqrt{\text{km}}$	Cyclic prefix	38 samples
PDL Γ_{dB}	0 – 6 dB	T_{OFDM}	25.7 ns
Non-linear coeff. γ_{NL}	$1.5 \text{ km}^{-1} \text{ W}^{-1}$	T_S	22.4 ns
EDFA noise figure	NF = 5 dB	Training sequence	32 OFDM symbols

in the “PT encoder” block to create new correlated symbols when a PT coding scheme is applied. The main simulation parameters are summarized in Table 3.1 and the simulation work is carried on MATLAB. The optical carrier is chosen at $\lambda_0 = 1550$ nm.

Each OFDM frame contains 500 data symbols preceded by a 32 symbol-long training sequence to estimate the channel using LSE estimation as explained in section 2.3.2.2. Among the 256 data subcarriers, 224 effective subcarriers are used for data symbols, and $N_p = 5$ pilot subcarriers for phase noise correction. The remaining 32 subcarriers are the outermost of the spectrum and are set to zero in order to prevent spectrum aliasing products. The cyclic prefix length was computed according to (2.12) in order to absorb both CD- and PMD-induced delays. The maximum budgeted DGD was chosen as three times the mean DGD: $DGD_{max} = 3\overline{\Delta\tau} = 3PMD\sqrt{L_{tot}} = 6.7$ to 134 ps for the parameters in table 3.1.

At the receiver, a polarization beam splitter splits the incoming signal on two orthogonal polarization states. Next, a dual polarization coherent receiver down-converts the optical signal to the electrical domain. Subsequently, a PDM OFDM receiver as the one in Fig. 2.5 is used to decode the data symbols. After proper time and frequency synchronizations and common phase error removal, each subcarrier sees a non-dispersive channel and the received symbols can be given by:

$$\mathbf{Y}_{k,2 \times T} = \exp(i\Phi_{CD}(\omega_k))\mathbf{H}_{k,2 \times 2}(\omega_k)\mathbf{X}_{k,2 \times T} + \mathbf{N}_{k,2 \times T} \quad (3.1)$$

where $\mathbf{X}_k$ (resp. $\mathbf{Y}_k$) is the matrix of transmitted (resp. received) symbols on the k^{th} subcarrier during T time slots, $\mathbf{N}_{k,i}$ the additive noise and $\mathbf{H}_k(\omega_k)$ is the 2×2 frequency-domain Jones matrix of the link. $\Phi_{CD}(\omega_k)$ is the phase shift owing to chromatic dispersion and is hence polarization independent. Non-linear effects in the fiber are so far neglected. T is the temporal length of the code. When $T = 1$, $\mathbf{X}$ and $\mathbf{Y}$ are vectors and $\mathbf{X}$ contains two independent q-QAM symbols (No Coding). When $T > 1$, PT coding is used ($T = 2$ for the investigated PT codes).

In section 2.4.4, we noted that the coding gains of ST codes are acquired when Maximum-Likelihood (ML) detection is performed at the receiver side. Hence, ML decoding will be also performed for PT codes in the PDM-OFDM optical system through an exhaustive search decoder. PT decoding can only be performed under the assumption of a constant $\mathbf{H}_k$ during the codeword duration (T time slots). In PDM optical transmissions, PMD dictates the coherence bandwidth of the channel that varies at a scale of few milliseconds due to environmental disturbances. Hence, the channel changes slowly in comparison to the codeword time that corresponds to two consecutive OFDM symbols ($2T_{OFDM} \sim 50$ ns in our case) and can be considered constant.

In the following, the channel model of the link $\mathbf{H}_k(\omega_k)$ will be described and simplified to carry the theoretical studies of PDL mitigation with PT codes. We will, in particular, look at the statistics of PDL, its frequency dependence and the noise properties at the receiver.

3.2.1 Performance of Silver, Golden and Alamouti coded schemes

To get an insight on the behavior of PT codes on an optical link with PDL, we consider three scenarios where we gradually add the following effects and study, for each, the PDL-induced penalties:

- First, we consider a long-haul multi-span channel with a *single lumped* PDL element as shown in Fig. 3.2a. Hence, we do not look at the interplay between PDL and PMD. Γ_{dB} is hence deterministic.
- Second, we distribute low-PDL elements in a single-span channel and add the ASE noise at the receiver as seen in Fig. 3.2b. In this case, Γ_{dB} becomes Maxwellian distributed because of the interaction of PDL and birefringent elements.
- Third, we consider a multi-span channel with distributed PDL as well as distributed ASE noise as shown in Fig. 3.2c in order to account for the polarization of distributed noise.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

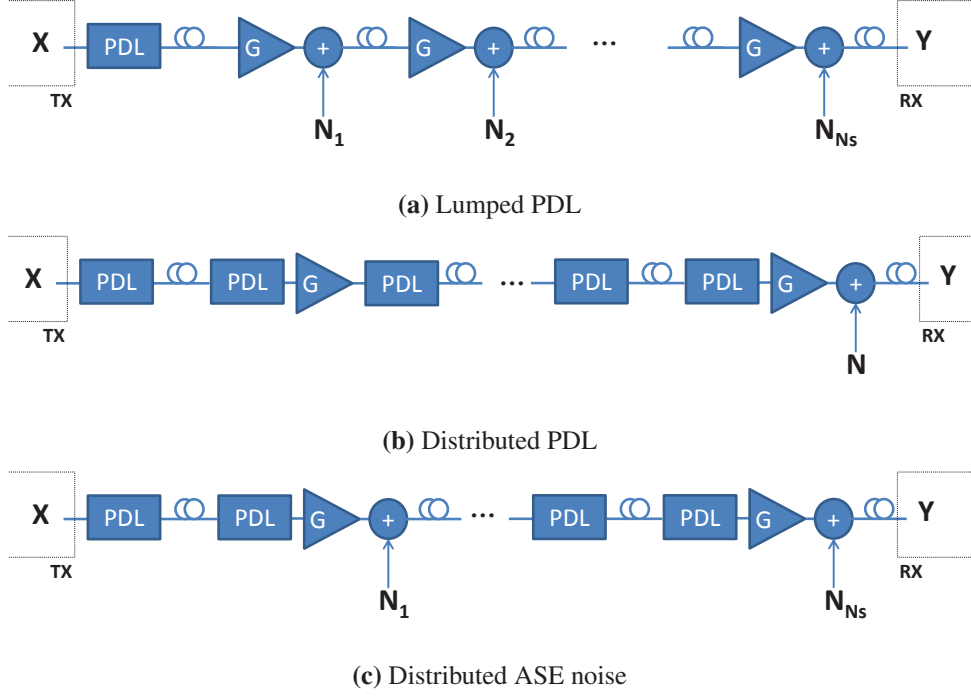


Figure 3.2: The three investigated long-haul optical links

3.2.1.1 Constant lumped PDL

We define a long-haul link of N_{spans} where each span is modeled as a concatenation of random PMD elements as in (1.19) followed by an optical amplifier operating at a constant output power to raise the signal power to its initial value at the transmitter. A single PDL element modeled by (1.21) is added in the first span with a constant $\Gamma_{dB} = 10 \log_{10} \frac{1+\gamma}{1-\gamma}$ and random rotation angles. The rotation angles of the different PMD and PDL elements are uniformly drawn from $[0 : 2\pi]$. In the PMD matrices, the phase shifts ϕ_i are uniformly drawn from $[0 : 2\pi]$ and the DGD τ_i are drawn from a normal distribution $N(0, \text{PMD}\sqrt{L})$ where PMD is the PMD coefficient $0.1 \text{ ps}/\sqrt{\text{km}}$ and $L = 1 \text{ km}$ is the length of a fiber slice. The distributed noise at each amplification stage is modeled as additive white with independent and identically distributed (iid) circular complex Gaussian components $CN(0, N_{ASE})$. In this case, the accumulated noise $\mathbf{N}$ is propagating through a unitary link, hence it remains white Gaussian with zero-mean and variance $N_0 = N_{spans}N_{ASE}/2$ per real dimension. The only source of power imbalance being the single PDL element, a singular value decomposition (SVD) of $\mathbf{H}$ yields two singular values $\lambda_{min} = \sqrt{1-\gamma}$ and $\lambda_{max} = \sqrt{1+\gamma}$ with a constant Γ_{dB} . Hence, (3.1) can be written as:

$$\begin{aligned}
 \mathbf{Y}_k &= \exp(i\Phi_{CD})\mathbf{H}_k(\omega_k)\mathbf{X}_k + \mathbf{N}_k \\
 &= \exp(i\Phi_{CD})\mathbf{U}_k(\omega_k)\mathbf{\Lambda}\mathbf{V}_k(\omega_k)\mathbf{X}_k + \mathbf{N}_k \\
 &= \exp(i\Phi_{CD})\mathbf{U}_k(\omega_k) \begin{bmatrix} \sqrt{1+\gamma} & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \mathbf{V}_k(\omega_k)\mathbf{X}_k + \mathbf{N}_k
 \end{aligned} \tag{3.2}$$

where $\mathbf{U}_k$ and $\mathbf{V}_k$ are two unitary matrices resulting from the SVD of the channel matrix.

We consider the OFDM system in Fig. 3.1 with 4-QAM symbols used for the uncoded case as well as to fill the codeword matrices of the Silver and the Golden codes. 16-QAM symbols are used for the Alamouti code in order to compare all schemes at the same rate of 4 bits/cu. Moreover, the average symbol energy E_S is set to 1 for all the schemes to compare them at the same power budget. No FEC coding is employed yet. The parameters of the optical link and the OFDM format are the ones presented in table 3.1. Monte-Carlo simulations were carried in order to evaluate the performance of the different coding schemes in terms of BER curves. The BER are averaged over all subcarriers and over the random angles of the rotation matrix. A minimum of 50 channel states are tested and a minimum of 100 bit errors recorded for each SNR value.

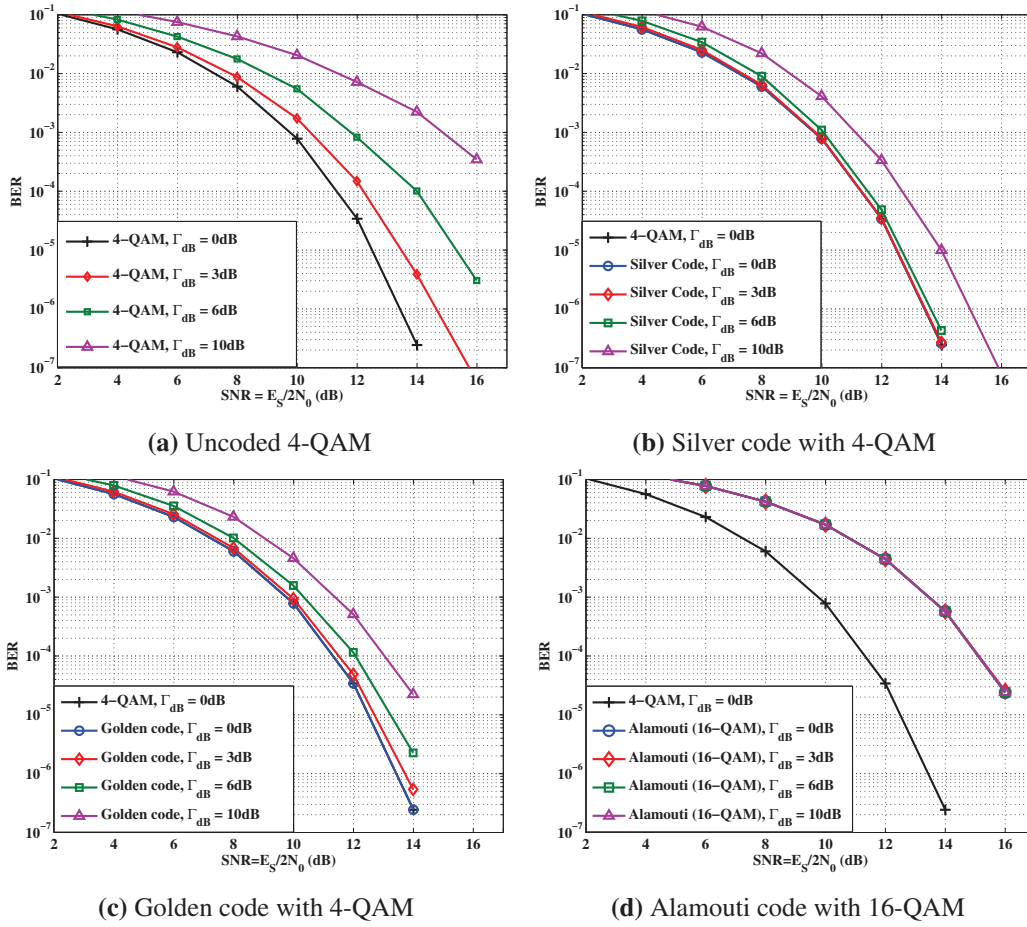


Figure 3.3: Average BER curves of 4-QAM (No coding), Silver, Golden and Alamouti coding schemes for different constant Γ_{dB} in a 20×100 km link.

Fig. 3.3 shows the BER evolution versus SNR for the four coding schemes at $\Gamma_{dB} = \{0, 3, 6, 10\}$. We can clearly see the efficiency of PT coding against PDL. Without PT codes, the PDL-induced SNR penalty at $\text{BER} = 10^{-3}$ (i.e. the gap to a perfect PDL-free AWGN channel) is 2.3 dB for a PDL of 6 dB. Using PT codes, the SNR penalty is only 0.6 dB with the Golden code and is

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

reduced to 0.3 dB with the Silver code corresponding to a coding gain of 2 dB. For a PDL of 3 dB, the Silver code absorbs all the penalty while a penalty of 1 dB is observed for a high PDL of 10 dB. On the other hand, the Alamouti code exhibits an SNR penalty of 3.8 dB at $\text{BER} = 10^{-3}$ for the three PDL levels which corresponds to the same penalty demonstrated on a PDL-free channel. This high value is due to the use of 16-QAM symbols that have a smaller minimum distance than 4-QAM symbols at the same average energy. However, an interesting observation is that Alamouti code performs the same independently of the PDL level. Had we considered the Alamouti code with 4-QAM symbols, the obtained performance would have matched the one over a PDL-free AWGN channel while the achieved spectral efficiency would have only been of 2 bits/cu.

These results coincide with initial ones reported in [8] where penalties were computed over a single subcarrier using the transfer matrix of a single PDL element $\mathbf{H}_{PDL}$, composed of a diagonal matrix with loss disparities and two real random rotation matrices, as the optical channel. While $\mathbf{H}_{PDL}$ and $\mathbf{H}_k \forall k$ have the same constant PDL values given that we intentionally kept PDL lumped in this long-haul channel model, the difference between these matrices is spotted at the level of the unitary transformations applied on the transmitted codewords $\mathbf{X}$. The SVD decomposition of $\mathbf{H}_k$ in (3.2) yields a complex unitary matrix $\mathbf{V}_k$ that can be written in its most general form as a product of three unitary matrices [169, Chap.2]:

$$\begin{aligned} \mathbf{V} &= \begin{bmatrix} \exp(i\alpha)\cos(\theta) & -\exp(i\beta)\sin(\theta) \\ \exp(-i\beta)\sin(\theta) & \exp(-i\alpha)\cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \exp(iu) & 0 \\ 0 & \exp(-iu) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \exp(iv) & 0 \\ 0 & \exp(-iv) \end{bmatrix} \end{aligned} \quad (3.3)$$

where $u = (\alpha + \beta)/2$ and $v = (\alpha - \beta)/2$. The unitary matrix $\mathbf{U}_k$ from the SVD decomposition of $\mathbf{H}_k$ has no influence on the performance as it can be inverted at the receiver without changing the properties of the noise $\mathbf{N}_k$. Hence, the similarities in the performance over this channel model and over $\mathbf{H}_{PDL}$ suggest that the angles α and β in $\mathbf{V}_k$ or equivalently u and v in the second form in (3.3) have no influence on the performance. This is evident for u because it appears in a diagonal matrix that can be moved to the other side of Λ and merged with $\mathbf{U}_k$. As for v , we carry out numerical simulations in which we vary θ and v and observe that the obtained performance in terms of BER, given in Fig. 3.4 at $\text{SNR} = 10$ dB, does not depend on the latter parameter, thus justifying the omission of v in a simplified model of an optical channel with PDL. Moreover, we notice that the penalties are periodic in θ with a period of $\pi/2$. For non-zero PDL, the BER penalty is maximal when the incoming signal is aligned to the principal polarization axes of the PDL element and minimal, but not null, when the mismatch angle is at $\pi/4 \pmod{\pi/2}$ radians.

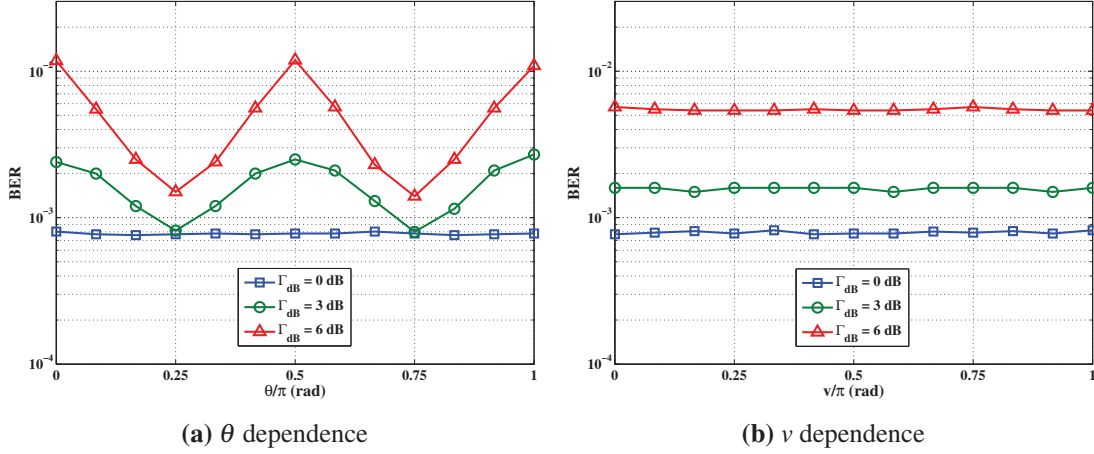


Figure 3.4: BER dependence on the parameters θ and ν of the unitary matrix $\mathbf{V}_k$ in the SVD decomposition of $\mathbf{H}_k$ for two levels of PDL at SNR = 10 dB.

3.2.1.2 Maxwellian distributed PDL

In the second model, we consider a multi-span link with N_{spans} spans where each fiber span is modeled as a concatenation of PMD and PDL elements. We divide each span into M random birefringent slices and add two PDL elements having i.i.d. PDL values, one placed at the start of the span and another at the end. Thus, the channel matrix of each span j is given by:

$$\begin{aligned} \mathcal{H}_j(\omega_k) &= \mathbf{H}_{PDL,2} \mathbf{H}_{PMD,M}(\omega_k) \mathbf{H}_{PMD,M-1}(\omega_k) \dots \\ &\quad \mathbf{H}_{PMD,2}(\omega_k) \mathbf{H}_{PMD,1}(\omega_k) \mathbf{H}_{PDL,1} \end{aligned} \quad (3.4)$$

In this model, ASE noise will be additive white Gaussian and lumped at the receiver with zero-mean and variance N_0 per real dimension as in the first model. Given that the optical amplifiers are operated in a constant output power mode, the received signal $\mathbf{Y}_k$ at the output of the link can be written in function of the input signal $\mathbf{X}_k$ and the lumped noise as:

$$\begin{aligned} \mathbf{Y}_k &= \exp(i\Phi_{CD}) \mathbf{H}_k(\omega_k) \mathbf{X}_k + \mathbf{N}_k \\ &= \exp(i\Phi_{CD}) \frac{\mathcal{H}_{1 \rightarrow N_{spans}} \sqrt{2}}{(\text{Tr}(\mathcal{H}_{1 \rightarrow N_{spans}} \mathcal{H}_{1 \rightarrow N_{spans}}^\dagger))^{1/2}} \mathbf{X}_k + \mathbf{N}_k \\ &= \exp(i\Phi_{CD}) \mathbf{U}_k(\omega_k) \mathbf{\Lambda}_k(\omega_k) \mathbf{V}_k(\omega_k) \mathbf{X}_k + \mathbf{N}_k \\ &= \exp(i\Phi_{CD}) \mathbf{U}_k(\omega_k) \begin{bmatrix} \sqrt{1 + \gamma_{eq,k}} & 0 \\ 0 & \sqrt{1 - \gamma_{eq,k}} \end{bmatrix} \mathbf{V}_k(\omega_k) \mathbf{X}_k + \mathbf{N}_k \end{aligned} \quad (3.5)$$

with $\mathcal{H}_{1 \rightarrow N_{spans}} = \mathcal{H}_{N_s} \mathcal{H}_{N_s-1} \dots \mathcal{H}_2 \mathcal{H}_1$. Normalizing the trace of the channel matrix to 2 depicts the constant output power mode of the amplifiers. An SVD of $\mathbf{H}_k$ yields in this case a diagonal matrix with frequency-dependent singular values $\lambda_{max} = \sqrt{1 + \gamma_{eq,k}}$ and $\lambda_{min} = \sqrt{1 - \gamma_{eq,k}}$, and hence a random accumulated PDL $\Gamma_{k,dB}$. We analyze these two new findings in this section.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

When the individual PDL values γ_i are i.i.d. (constant, Gaussian distributed or uniformly distributed; by the law of large numbers, the nature of the distribution is no longer important when the number of PDL elements is sufficiently large (> 20)), it is found that the total Γ_{dB} of the link is Maxwellian distributed [53] as noted in section 1.3.4. The accuracy of this approximation is tested through a numerical simulation. We consider 40 PDL elements and target an average overall PDL $\bar{\Gamma}_{dB}$ of 3 and 6 dB. The average PDL fixes the value of the parameter a of the Maxwellian distribution in (1.23) to 1.88 and 3.76 respectively. Following the procedure in [53], each PDL element should have a γ_i drawn from a Gaussian distribution of 0.047 (resp. 0.092) and standard deviation 0.035 (resp. 0.069) in order to achieve an overall average PDL of 3 dB (resp. 6 dB). The parameters in the transfer matrices of the PMD elements and the rotation matrices of the PDL elements are chosen as in the first model. We simulate 10^6 different realizations of the link in (3.5) and show, in Fig. 3.5, the obtained probability distribution of $\Gamma_{k,dB}$ at any given subcarrier k . We can see that the fitting to a Maxwellian distribution is accurate for both average PDL.

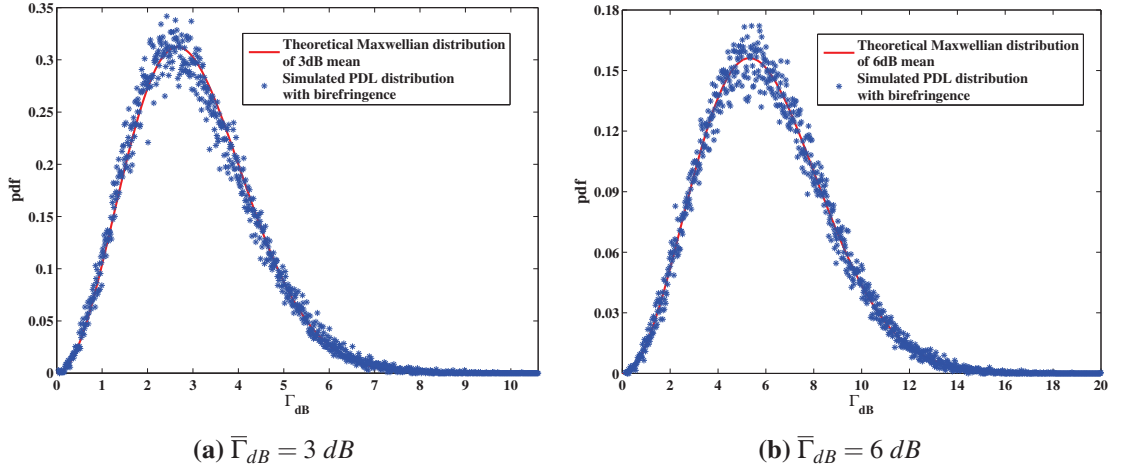


Figure 3.5: Simulated and theoretical probability distribution functions of the overall PDL Γ_{dB} of a concatenation of low-PDL and birefringent elements.

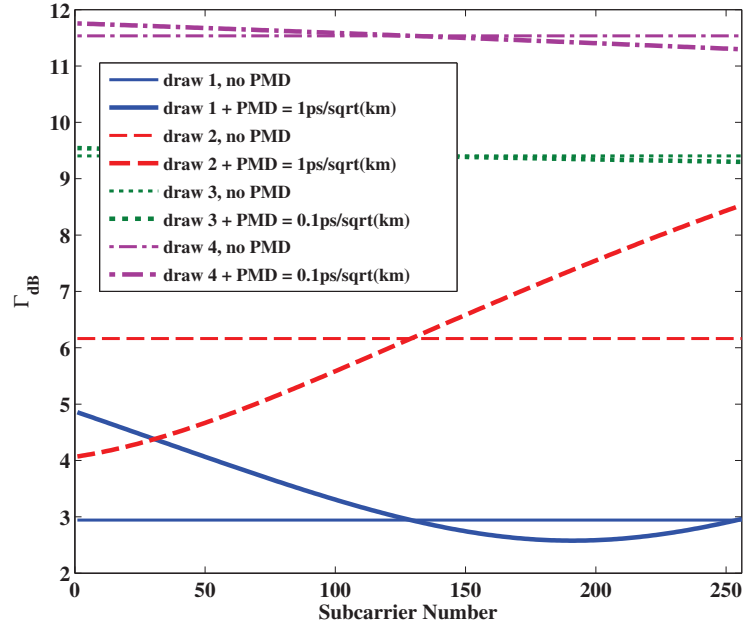
We also performed simulations where γ_i were constants or uniformly distributed and ended up with the same Maxwellian distributions of the overall PDL. The PDL at the end of the transmission link is therefore statistically distributed, and does not linearly increase with the number of PDL components in the link. Moreover, in order to determine the effect of birefringence on the distribution of $\Gamma_{k,dB}$, we ran identical simulations with a null PMD coefficient and also obtained a Maxwellian distribution. However, if we remove the phase shifts ϕ_i from the PMD matrices, hence totally removing birefringence in the fiber, the simulated $\Gamma_{k,dB}$ fits a Rayleigh distribution with the same average Γ_{dB} . This result was noticed by Lu *et al.* in [170] where the authors also suggested a global distribution made of a linear combination of a Maxwellian and Rayleigh distributions depending on the amount of birefringence in the fiber.

On the other hand, (3.5) shows that the channel matrix is frequency dependent. This might induce fluctuations of the Γ_{dB} experienced by the different OFDM subcarriers. We investigate into these variations through a comparison of 10^6 realizations of a multi-span optical link with $\bar{\Gamma}_{dB} = 6$ dB for two values of the PMD coefficient: $\text{PMD} = \{0.1, 1\}$ ps/ $\sqrt{\text{km}}$. For the first value, we noticed that the deviations of Γ_{dB} of the different subcarriers from its constant value in a PMD-free case are very small. Averaging over 10^6 realizations and all subcarriers, we obtain $\mathbb{E}[|\Delta\Gamma_{dB}|] = 0.09$ dB and $\text{Var}[|\Delta\Gamma_{dB}|] = 0.01$ dB. In the case of highly birefringent fibers with $\text{PMD} = 1$ ps/ $\sqrt{\text{km}}$ (old fibers deployed before 90s), the deviations of Γ_{dB} were more important. This can be seen in Fig. 3.6a where Γ_{dB} is shown at the different subcarriers for two random channel realizations. A channel realization is randomly drawn and the PDL is observed in absence (thin curves) and in presence (thick curves) of PMD for two different PMD coefficients. The frequency dependence of Γ_{dB} due to PMD was reported in [171]. To generalize these findings, the frequency variations of Γ_{dB} over the signal bandwidth will be determined by the coherence bandwidth of the channel B_{coh} dictated by PMD ($B_{coh} \propto \overline{DGD}^{-1}$ [58]). If the total bandwidth of the OFDM signal $\approx N_{sc}\Delta f_S$ is smaller than B_{coh} , all the subcarriers will observe the same PDL. Otherwise, PDL will vary among subcarriers.

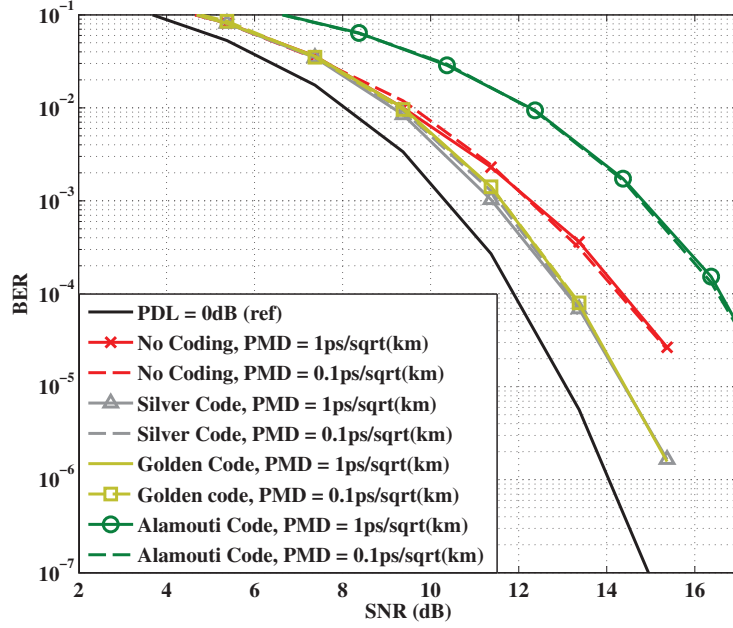
The average BER performance of the PDM-OFDM transmission for two different PMD coefficients corresponding to a frequency-varying and a frequency-flat PDL is shown in Fig. 3.6b with a mean PDL of 6 dB. We ran Monte-Carlo simulations of the PT-coded PDM-OFDM system with this second channel model using the parameters in table 3.1 and noticed that the average performance is insensitive to the frequency variation. The frequency dependence of Γ_{dB} due to PMD was reported in [133, 172]. In [133], the authors found that an increased DGD (up to a mean DGD of 100 ps) in the optical link reduced the BER variations caused by PDL and did not alter the mean BER value.

Again, we observe the efficiency of PT coding against PDL even when Γ_{dB} is random and follows a Maxwellian distribution. The SNR penalty at $\text{BER} = 10^{-3}$ is only 1.3 dB with the Golden code and is reduced to 1 dB with the Silver code corresponding to a coding gain of 1 dB. On the other hand, the Alamouti code still exhibits a PDL-independent SNR penalty of 3.8 dB due to the use of 16-QAM symbols. The behaviors of PT codes are unchanged on this second channel model while the coding gain values are different because they are averaged over the random Maxwellian distributed PDL. Indeed, for each channel realization, a different value of PDL is achieved. The Silver code may experience a PDL less than 6 dB for a certain realization and absorbs all the penalty in this case; and then experience a larger PDL for another random draw for which an SNR penalty still remains. The same observation can be made for the Golden code. However, the Alamouti code still performs the same independently of Γ_{dB} . The same results are obtained for a 3 dB-mean Maxwellian PDL.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS



(a) Four random draws showing the evolution of $\Gamma_{k,dB}$ among the OFDM subcarriers for two different PMD coefficients and $\bar{\Gamma}_{dB} = 6$ dB.



(b) Average BER performance of a 4-QAM PDM-OFDM transmission for two different PMD coefficients and $\bar{\Gamma}_{dB} = 6$ dB.

Figure 3.6: Frequency-dependence of PDL due to PMD in the optical fiber link.

3.2.1.3 Distributed ASE noise

The predominant noise in long-haul optical transmissions is the amplified spontaneous emission noise at the optical amplifiers, modeled as an additive white Gaussian noise. However, the distributed PDL elements act as a partial polarizer affecting both the signal and the in-line inserted noise propagating through the link. Given that the optical amplifiers are operated in a constant output power mode, the received signal $\mathbf{Y}_k$ at the output of the third link model in Fig. 3.2c can be written in function of the input signal $\mathbf{X}_k$ and the in-line injected noise as:

$$\begin{aligned} \mathbf{Y}_k &= \underbrace{\frac{\mathcal{H}_{1 \rightarrow N_{spans}} \sqrt{2}}{(\text{Tr}(\mathcal{H}_{1 \rightarrow N_{spans}} \mathcal{H}_{1 \rightarrow N_{spans}}^\dagger))^{1/2}}}_{\mathbf{H}_k} \mathbf{X}_k + \underbrace{\sum_{j=1}^{N_{spans}-1} \frac{\mathcal{H}_{j+1 \rightarrow N_{spans}} \sqrt{2}}{(\text{Tr}(\mathcal{H}_{j+1 \rightarrow N_{spans}} \mathcal{H}_{j+1 \rightarrow N_{spans}}^\dagger))^{1/2}} \mathbf{N}_j}_{\mathcal{N}_k} + \mathbf{N}_{N_s} \\ &= \mathbf{H}_k \mathbf{X}_k + \mathcal{N}_k \end{aligned} \quad (3.6)$$

with $\mathcal{H}_{j \rightarrow N_{spans}} = \mathcal{H}_{N_s} \mathcal{H}_{N_s-1} \dots \mathcal{H}_{j+1} \mathcal{H}_j$. Each component of the noise vector $\mathbf{N}_j$ added after the j^{th} span is white and Gaussian distributed with a zero mean and a variance $N_{ASE}/2$ per polarization and per real dimension. Because of distributed PDL, $\mathcal{N}_k$ will be polarized and its coherency (or covariance) matrix $\mathbf{Q} = \mathbb{E}[\mathcal{N}_k \mathcal{N}_k^\dagger]$ will not be proportional to the identity matrix as is the case of a white Gaussian noise. The $\mathbb{E}[\cdot]$ operation in the coherency matrix is the temporal averaging of the considered quantity over an OFDM symbol period [169, chap.1]. However, we can still define a noise-whitened system given by:

$$\mathbf{Y}'_k = \mathbf{Q}^{-1/2} \mathbf{Y}_k = \mathbf{H}'_k \mathbf{X}_k + \mathbf{N}_k \quad (3.7)$$

where $\mathbf{N}_k$ is a zero-mean white Gaussian noise of variance N_0 per real dimension. The effective PDL experienced by the signal in this new system and given by the ratio of the eigenvalues of $\mathbf{H}'_k \mathbf{H}'_k{}^\dagger$ will hence be different from the PDL given by the ratio of the eigenvalues of $\mathcal{H}_k \mathcal{H}_k^\dagger$. We simulate 10^6 channel realizations of the third system model with the same parameters in table 3.1 fixing a mean PDL of 6 dB. The probability distribution function of the effective PDL is shown in Fig. 3.7. It can be fitted to a Maxwellian distribution of mean 4 dB. We notice that the mean effective PDL of $\mathbf{H}'_k$ is 2 dB lower than the mean PDL of $\mathbf{H}_k$. The smaller effective PDL values can be explained by the fact that the noise is also polarized by PDL and some of the orthogonal noise to each of the data-carrying polarization tributaries may become parallel due to the distributed PDL elements, which was observed in [59, 60]. In [60], the author compares a lumped PDL model to a distributed PDL model and concludes that the former overestimates the PDL-induced penalties.

Hence, we can carry out the analysis of a PDL-impaired optical system using the second model with a lumped white Gaussian noise added at the receiver and keep in mind that the effective PDL observed by the signal is less than the measured PDL due to the polarization of noise.

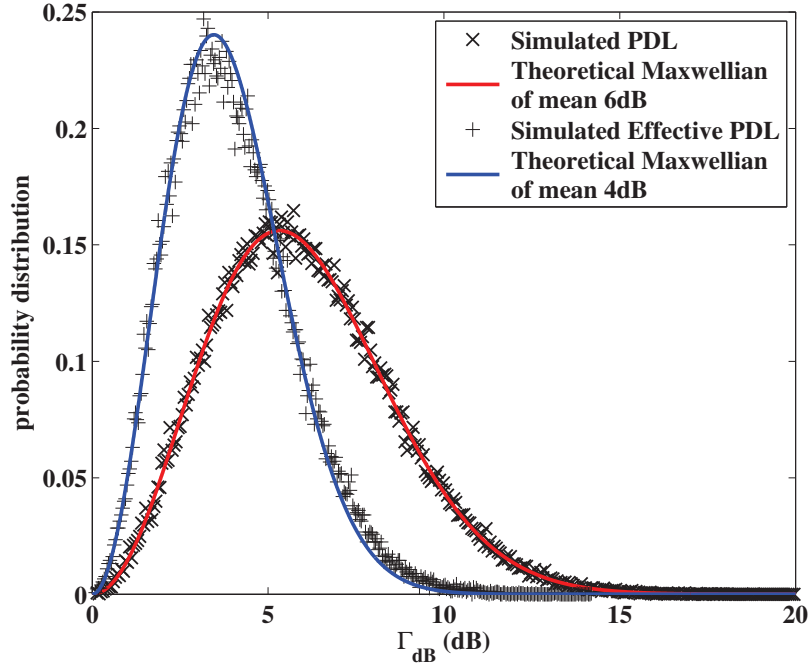


Figure 3.7: In ‘×’ marks, the simulated distribution of Γ_{dB} and the corresponding theoretical Maxwellian distribution of mean 6 dB in red. In ‘+’ marks, the simulated distribution of the effective PDL of $\mathbf{H}_k$ and the corresponding theoretical Maxwellian of mean 4 dB in blue.

Moreover, knowing that the dynamics of the optical channel present rather slow variations compared to the PT codeword duration, we will only regard constant values of global Γ_{dB} , corresponding to the largest achieved PDL value $20 \log_{10} \frac{1+\gamma_{eq}}{1-\gamma_{eq}}$ (worst-case value), when going through the theoretical analysis of the coding gains. Besides, from the simulation of the first model, we have seen that the BER performance only depended on the angle of the rotation matrix in the unitary matrix $\mathbf{V}_k$ derived from the singular-value decomposition of the channel and applied to the signal $\mathbf{X}_k$ before the diagonal matrix Λ_k containing the gain offsets. Hence, to theoretically interpret the performance of PT codes over PDL-impaired optical channels, the following frequency-flat channel matrix will be sufficient:

$$\mathbf{H} = \begin{bmatrix} \sqrt{1+\gamma_{eq}} & 0 \\ 0 & \sqrt{1-\gamma_{eq}} \end{bmatrix} \mathbf{R}_{\theta} \quad (3.8)$$

where $\mathbf{R}_{\theta}$ is a real rotation matrix and θ is drawn from a uniform distribution in $[0 : 2\pi]$. Before providing the theoretical analysis of the coding gains and checking whether the Silver code is the optimal code for an optical channel with PDL or better codes can be found, we still need to test the impact of non-linear effects on PT-coded OFDM systems.

3.2.2 Tolerance of PT-coded OFDM for non-linear effects

In this section, we evaluate the performance of the PT-coded PDM-OFDM system taking into account non-linear transmission effects. Indeed, the fiber is a non-linear medium having a refractive index that changes proportionally to the square of the applied electromagnetic field, an effect known as the optical Kerr effect as explained in chapter 1. Numerical solutions for the non-linear partial differential equation in (1.7) are computed using the Split-Step Fourier method [19, chap.2]. This method consists of partitioning the fiber into short consecutive dispersion-only and pure non-linear sections. The dispersion-only sections are modeled in the frequency domain and contain the transfer matrices of CD and PMD. The pure non-linear sections are modeled in the time domain. The spatial step size Δz is chosen small enough to correctly model both dispersion and non-linear effects. The simulation model implemented in MATLAB is shown in Fig 3.8.

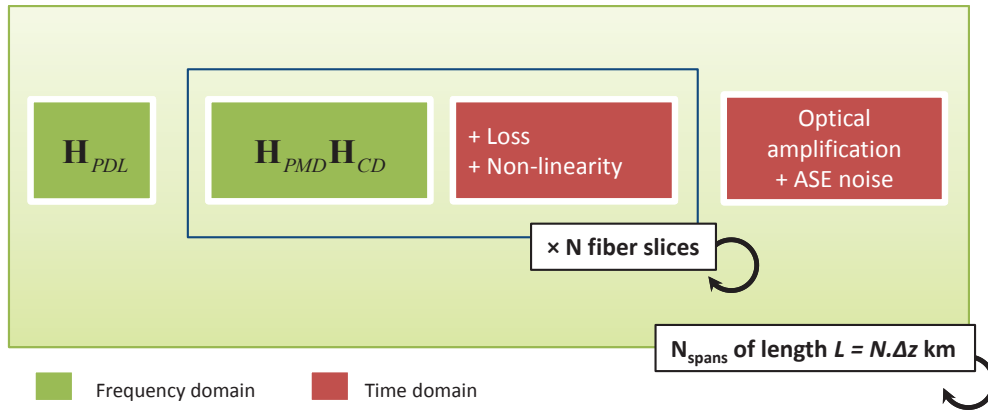


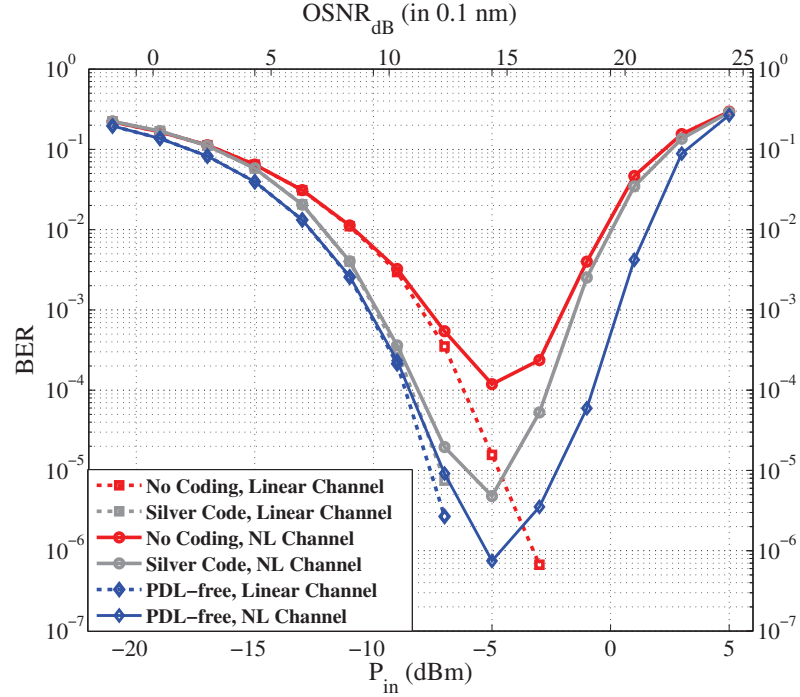
Figure 3.8: Simulation model of a long-haul PDM optical fiber link using the Split-Step Fourier method to include non-linear effects in the SMF fiber.

The simulation parameters are the one shown in table 3.1. Apart from dispersion and attenuation, a lumped PDL element as in (1.21) is inserted in the first span. We consider constant Γ_{dB} values ranging from 0 to 6 dB. Initially, spatial step sizes Δz from 0.2 to 5 km were tested and a final step size of 1 km was adopted because all smaller step sizes ensured the same numerical accuracy. At the receiver, we perform the OFDM de-multiplexing followed by channel estimation, CPE correction and data detection using an ML decoder.

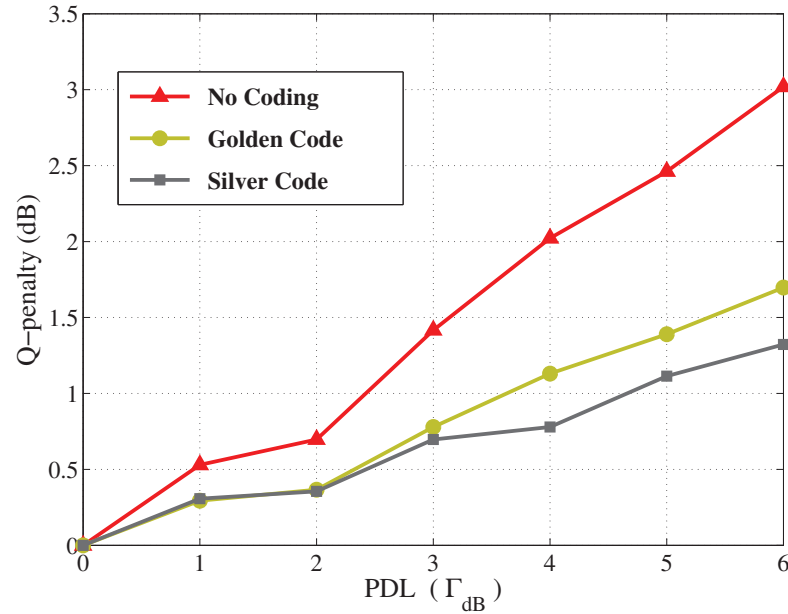
In Fig. 3.9a, we compare the performance of the Silver code to that of the uncoded scheme in the case of a linear and non-linear channel for a PDL of 6 dB. When non-linearity is considered, BER curves are bathtub-shaped and we can distinguish 3 regimes: a weak, a medium and a severe non-linear regime. Below a certain input power threshold, at -7 dBm, the obtained performance perfectly matches the curve of the linear model. Then, if we increase the input power to -3 dBm, we observe a transient behavior of the curves where a minimum BER is reached (the non-linear threshold NLT). Beyond, we enter a severe non-linear regime where BERs are totally deteriorated

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

for both schemes. By comparing the coded and uncoded cases, we clearly see the efficiency of the Silver code in mitigating PDL. A 2.2 dB coding gain is observed at $\text{BER} = 10^{-3}$. This value matches the gain in the previous section and in [8, 167].



(a) BER evolution of Silver-coded and uncoded scenarios versus total launched power after a 2000 km transmission (Lumped PDL = 6 dB).



(b) Q -factor penalty induced by PDL after 2000 km (with regard to an uncoded PDL-free scheme).

Figure 3.9: Performance of PT-coded PDM-OFDM system in presence of Kerr non-linear effects.

Apart from the horizontal coding gains at a fixed BER, we can evaluate vertical BER enhancement at a given input power. At an input power of -5 dBm, corresponding to a total OSNR of 14 dB, an optimal BER = $5 \cdot 10^{-6}$ is achieved with the Silver code compared to a BER = 10^{-4} for the uncoded case. Furthermore, the Silver code does not add any extra penalty in the severe non-linear regime compared to the uncoded case (the same behavior is also observed with the Golden code). This is mainly due to the fact that both coding schemes preserve the average energy of the symbols sent on each polarization.

In Fig. 3.9b, we show the Q -factor penalties at the NLT of -5 dBm for the Golden-, Silver- and uncoded schemes at different PDL values. Optimal BER and Q -factors are related by:

$$\begin{aligned} \text{BER} &= \frac{1}{2} \text{erfc}(Q/\sqrt{2}) \\ Q_{\text{dB}} &= 20 \log_{10}(Q) \end{aligned} \tag{3.9}$$

where $\text{erfc}()$ is the complementary error function defined as: $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dx$. The penalties are computed with regard to an uncoded scheme and a PDL-free link. We notice that penalties are smaller when PT coding is used. A 3 dB Q -factor penalty is obtained without coding for a PDL = 6 dB, whereas the penalty is only of 1.3 dB with the Silver code.

In conclusion, non-linear effects induce no extra penalty when PT coding with a uniform average energy per polarization is used, and the coding gains are maintained in the weak and medium non-linear regimes around the NLT.

3.3 Theoretical analysis of the PT coding performance

The suggested PT codes were designed for wireless communications and their performance on an optical channel with PDL showed a different behavior leading to important questions: is the Silver code the optimal for an optical channel? Or can we find better coding schemes?

In order to answer these questions, we will follow the same analysis that has been carried out to design codes for the wireless channel. Some previous works have investigated the performance of PDM schemes in presence of PDL in terms of the outage probability (probability that the transmitted bitrate is greater than the actual channel capacity) [173] or the system outage defined as the probability that the SNR penalty exceeds a given threshold [59, 166]. However, in both cases, outage analyses only give a behavioral description of the transmission schemes without any explicit implication of the transmitted coding scheme. The only information related to the transmitted signal that appears in an outage analysis is the symbol energy and the transmission rate. On the other hand, an error probability analysis based on the calculus of an upper bound of the probability that an ML decoder erroneously decodes the transmitted symbol takes into account the coding scheme

and is the only way to extract the properties that the codeword matrix should have in order to mitigate the channel impairments.

An error probability calculus in the case of a MIMO Rayleigh fading channel can be found in [140]. This calculus revealed the design criteria used to construct the Golden and the Silver codes for the wireless channel. In the following, we derive an upper bound of the pairwise error probability (PEP) and design criteria of PT codes in the case of an optical channel with PDL.

3.3.1 Computing an upper bound of the pairwise error probability (PEP)

An upper bound of an error probability expression is found by averaging over all possible realizations of the channel as explained in section 2.4.2. To get an insight on the behavior of PT codes on an optical link with PDL, we consider a PDM-OFDM transmission and use $\mathbf{H}$ defined in (3.8) as the channel matrix. Each subcarrier is seen as a block-fading (the optical channel varies among a block of hundreds or thousands of OFDM symbols) and frequency non-selective channel. We have seen in 2.4.2 that the PEP of an ML-decoded MIMO channel is given by:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \mathbb{E}_H \left[\exp \left(- \frac{\|\mathbf{H}(\hat{\mathbf{X}} - \mathbf{X})\|^2}{8N_0} \right) \right] \quad (3.10)$$

where $\mathbb{E}_H[\cdot]$ is the averaging operation over all possible channel realizations, $\mathbf{X}$ the transmitted codeword and $\hat{\mathbf{X}}$ the estimated one. In this case, the only random variable in $\mathbf{H}$ is the random rotation angle θ .

Let $\mathbf{X}_\Delta = \hat{\mathbf{X}} - \mathbf{X}$ be the difference of two codewords and $\Lambda = \begin{bmatrix} \sqrt{1-\gamma_{eq}} & 0 \\ 0 & \sqrt{1+\gamma_{eq}} \end{bmatrix}$, we have:

$$\|\mathbf{H}(\hat{\mathbf{X}} - \mathbf{X})\|^2 = \|\mathbf{H}\mathbf{X}_\Delta\|^2 = \|\Lambda\mathbf{R}_\theta\mathbf{X}_\Delta\|^2 \quad (3.11)$$

We can write $\mathbf{R}_\theta\mathbf{X}_\Delta = \mathbf{Z} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}$, $\mathbf{z}_{1,2}$ being row vectors and $\mathbf{R}_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$, we get:

$$\|\Lambda\mathbf{Z}\|^2 = \|\mathbf{z}_1\|^2 + \|\mathbf{z}_2\|^2 + \gamma_{eq}(\|\mathbf{z}_2\|^2 - \|\mathbf{z}_1\|^2) \quad (3.12)$$

However, $\|\mathbf{z}_1\|^2 + \|\mathbf{z}_2\|^2 = \|\mathbf{Z}\|^2 = \|\mathbf{X}_\Delta\|^2$ because the rotation matrix is unitary. Besides, we develop $\|\mathbf{z}_2\|^2 - \|\mathbf{z}_1\|^2$ and obtain:

$$\begin{aligned} \|\mathbf{z}_2\|^2 - \|\mathbf{z}_1\|^2 &= f(\theta) = (\|\mathbf{x}_2\|^2 - \|\mathbf{x}_1\|^2) \cos(2\theta) \\ &\quad + 2\Re(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle) \sin(2\theta) \end{aligned} \quad (3.13)$$

with $\mathbf{X}_\Delta = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$, $\mathbf{x}_{1,2}$ being row vectors and the scalar product defined as $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \mathbf{x}_1 \mathbf{x}_2^\dagger$. Let:

$$a = \|\mathbf{x}_2\|^2 - \|\mathbf{x}_1\|^2 \quad (3.14)$$

$$b = 2\Re(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle) \quad (3.15)$$

Substituting (3.12) to (3.15) in (3.10), we get:

$$\begin{aligned} \Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) &= \mathbb{E}_H \left[\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}} | \mathbf{H}) \right] \\ &\leq \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) \mathbb{E}_H \left[\exp\left(-\frac{\gamma_{eq}}{8N_0}(a \cos(2\theta) + b \sin(2\theta))\right) \right] \end{aligned} \quad (3.16)$$

The last step consists in averaging over $\mathbf{H}$. Considering constant values of Γ_{dB} , we are left with one random variable θ that varies uniformly in $[0 : 2\pi]$. Using the following trigonometric identity:

$$\begin{aligned} a \cos(2\theta) + b \sin(2\theta) &= \sqrt{a^2 + b^2} \cos(2\theta + \phi) \\ \text{with } \phi &= \begin{cases} -\arctan(\frac{b}{a}) & \text{if } a > 0 \\ \pi - \arctan(\frac{b}{a}) & \text{if } a < 0 \end{cases} \end{aligned} \quad (3.17)$$

We obtain:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) \frac{1}{2\pi} \int_{\phi}^{2\pi+\phi} \exp\left(-\frac{\gamma_{eq}}{8N_0} \sqrt{a^2 + b^2} \cos(u)\right) du \quad (3.18)$$

The function inside the integral is periodic of period 2π and has an axis of symmetry $u = \pi$ in the interval $[0 : 2\pi]$. Using these two properties, (3.18) becomes:

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) \frac{1}{\pi} \int_0^\pi \exp\left(-\frac{\gamma_{eq}}{8N_0} \sqrt{a^2 + b^2} \cos(u)\right) du \quad (3.19)$$

Hence,

$$\Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) I_0\left(\frac{\gamma_{eq}}{8N_0} \sqrt{a^2 + b^2}\right) \quad (3.20)$$

where $I_0(k)$ is the 0^{th} order modified Bessel function of the first kind depicted in Fig. 3.10.

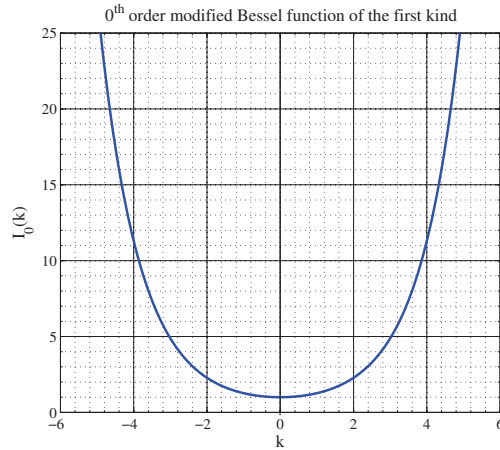


Figure 3.10: Plot of the 0^{th} order modified Bessel function of the first kind.

We can approximate the error probability expression in (3.20) for high SNR values by using a

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

first order approximation of $I_0(k)$ when $k \rightarrow \infty$ as follows:

$$I_0(k) \approx \frac{\exp(k)}{\sqrt{2\pi k}} \quad (3.21)$$

Then, we get:

$$\exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) I_0\left(\frac{\gamma_{eq}}{8N_0} \sqrt{a^2 + b^2}\right) \approx \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2 - \gamma_{eq} \sqrt{a^2 + b^2}}{8N_0}\right) \quad (3.22)$$

3.3.1.1 Design criterion of PT codes

In Fig. 3.10, we see that $I_0(k)$ is monotonously increasing for $k \geq 0$ and hence has its minimum for $k = 0$. This corresponds to a and b in (3.20) equal to zero and as a result, the obtained error probability expression will be independent of PDL. Consequently, we define the following design criterion:

Proposition 1. *A Polarization-Time code completely mitigates PDL if and only if all codeword differences satisfy:*

1. $a = \|\mathbf{x}_2\|^2 - \|\mathbf{x}_1\|^2 = 0$ and
2. $b = 2\Re(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle) = 0$.

The second condition in proposition 1 can be reinterpreted as an orthogonality between the equivalent real-valued vectors $\mathbf{x}_{R1}$ and $\mathbf{x}_{R2}$ where $\mathbf{x}_R = [\Re(\mathbf{x}) \ \Im(\mathbf{x})]$. When the design criterion is met, we recover the performance over two parallel additive white Gaussian channels which is the best performance that we can achieve and is given by:

$$p_{error, AWGN} \approx \exp\left(-\frac{\|\mathbf{x}_\Delta\|^2}{8N_0}\right) \quad (3.23)$$

We notice that the resulting criterion is completely different from the rank and the minimum determinant criteria derived for a Rayleigh fading channel [140] and defining respectively a diversity gain and a coding gain as explained in section 2.4.2. If we compare the approximation of the error probability expression at high SNR in (3.22) to the one obtained in the case of a 2×2 MIMO Rayleigh fading channel in (2.32), we notice different behaviors: the error probability of the Rayleigh fading channel decays as SNR^{-2r} [140] where r is the minimum rank of the matrix $\mathbf{X}_\Delta$ and the diversity order is $2r$ whereas the error probability of the PDL channel drops exponentially as a function of the SNR as in the case of an additive white Gaussian channel. Hence, Space-Time codes, used as Polarization-Time codes, bring no diversity gain to the optical channel with PDL (the exponential decrease of the error probability depicts an infinite diversity order). A coding gain that will be evaluated in the following, is only brought, hence reducing the penalty induced by PDL.

3.3.1.2 Code performance analysis

The performance of both coded and uncoded schemes will be examined using the derived upper bound of the PEP in (3.22). To compare the different schemes, we will compute the squared distance $d^2 = \|\mathbf{X}_\Delta\|^2 - \gamma_{eq}\sqrt{a^2 + b^2}$ of all combinations of codeword differences and compare the minimum values of d^2 , denoted d_{min}^2 , of the codes. Even though the actual upper bound of the error probability will depend on the sum of different PEPs as given by (2.27), the PEP with the smallest distance d^2 will be the dominant term and limit the performance. The best code is the one that maximizes d_{min}^2 . In Table 3.2, we report the minimum values of $\|\mathbf{X}_\Delta\|^2 - \gamma_{eq}\sqrt{a^2 + b^2}$ analytically computed for each investigated PT code at four different PDL values. To fill the table, we set a spectral efficiency of 4 bits/cu for all coding schemes and consider an average symbol energy $E_s = 1$ for all constellations.

Table 3.2: d_{min}^2 for different coding schemes at different PDL values

	$\Gamma = 0$ dB	$\Gamma = 3$ dB	$\Gamma = 6$ dB	$\Gamma = 10$ dB
Silver Code with 4-QAM	2	2	2	1.23
Golden Code with 4-QAM	2	1.7	1.46	1.07
No Coding with 4-QAM	2	1.34	0.8	0.38
Alamouti Code with 16-QAM	0.8	0.8	0.8	0.8

First, we notice that the Alamouti code has the same minimal distance for all PDL values which explains why it performs the same independently of PDL in Fig. 3.3d. This is due to the orthogonality of its codeword matrix (2.38) that induces $a = b = 0$ for all possible codeword differences. Hence, this code satisfies the criterion of *Proposition 1*. However, its performance is affected by the use of 16-QAM symbols giving a squared minimal distance of 0.8.

Second, we note that the Silver code is not optimal in mitigating PDL since it does not satisfy the derived design criterion. Unlike the Alamouti code, the Silver code has only some codeword differences having $a = b = 0$. The computed d_{min}^2 in Table 3.2 explains the behavior of the Silver code in Fig. 3.3b. d_{min}^2 is given by the same codeword difference with $a = 0$ and $b = 0$ at a PDL of 3 dB and 6 dB, and is equal to 2. Whereas at a PDL of 10 dB, it falls to 1.23 given by another codeword difference with $a \neq 0$ and $b \neq 0$.

Third, in Fig. 3.3, we saw that the Silver code outperforms the Golden code for $\Gamma_{dB} = 6$ dB, and both reduce the penalty that PDL causes to the uncoded scheme. Again, this result can be explained by looking at Table 3.2. Indeed, d_{min}^2 is the greatest for the Silver code followed by the Golden code and then the uncoded scheme. In conclusion, we were able to explain, in terms of error probability bounds, the performance of the Alamouti, Silver and Golden codes on an optical channel with PDL. These codes were designed to satisfy the rank and the minimum determinant criteria for a wireless channel that are no more relevant for the optical channel.

3.3.2 Concatenation of a FEC code and a PT code

Until now, we did not introduce any FEC coding into the transmission scheme. While PT coding uses the modulated symbols to form a codeword matrix and mitigate differential non-unitary effects in the MIMO channel, FEC or channel codes operate on the information bits and add some redundancy in order to enhance the performance over a noisy channel as explained in section 2.1. The FEC block at the transmitter side precedes the q-QAM modulation block and a corresponding FEC decoding unit at the receiver side follows the demodulation block as represented in Fig. 3.1. Given that the FEC and PT codes serve different purposes, we are interested in evaluating the total gain provided by FEC coding and PT coding through Monte-Carlo simulations.

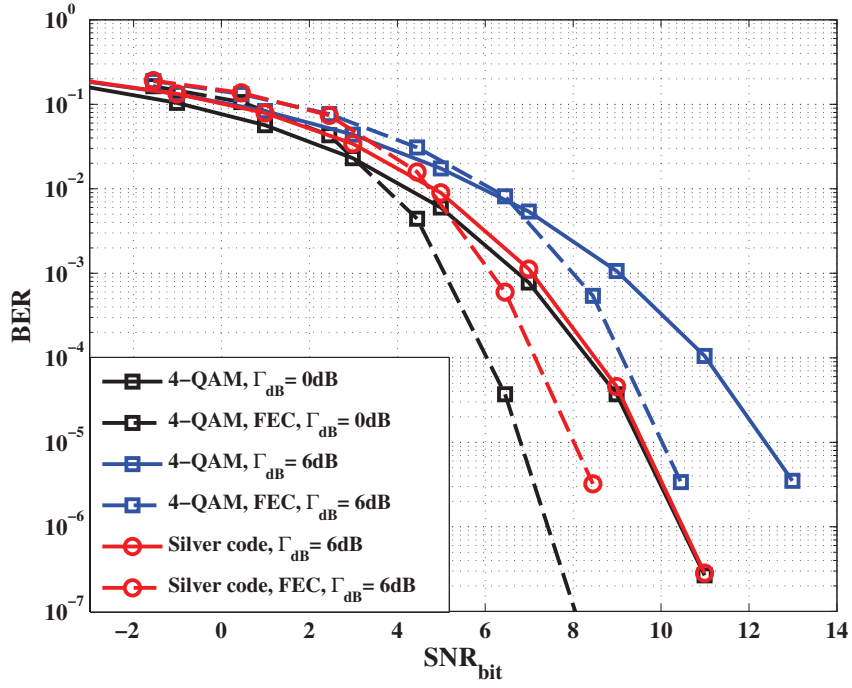
The performance over an optical channel with $\Gamma_{dB} = 6$ dB of an uncoded 4-QAM (no FEC, no PT) and a Silver-coded schemes are compared with a Bose-Chaudhuri-Hocquenghem (BCH) coded with $n = 63$, $k = 45$ and an error correction capability $t = 3$, and a BCH-coded and Silver-coded schemes. When FEC coding is considered, two information bit streams are independently encoded by the same FEC then assigned to the 4-QAM modulators. The obtained BER curves are plotted in Fig. 3.11. At a BER = 10^{-4} , the gain provided by the Silver code alone is 2.4 dB. When hard decision decoding is considered (performed by using a built-in MATLAB function based on the Berlekamp-Massey algorithm), the gain provided by the BCH code alone is 1.8 dB. The concatenation of both codes offers a total coding gain of 3.9 dB, approximately equal to the sum of the separate coding gains $2.4 + 1.8 = 4.2$ dB. The summation of the FEC and PT coding gains is also observed when using soft decision decoding through a Chase II algorithm [174]. In this case, the gain offered by the BCH code alone at a BER = 10^{-4} is 2.8 dB. The concatenation with the Silver code provides a total gain of 4.9 dB and the sum of the separate coding gains is $2.4 + 2.8 = 5.2$ dB.

We will also theoretically explain the summation of the separate coding gains brought by FEC and PT codes that is mainly due to the exponential decrease of the error probability as a function of SNR (Gaussian-like behavior leading to a linear summation of gains on a logarithmic scale). Hereafter, we show the theoretical proof for linear block codes.

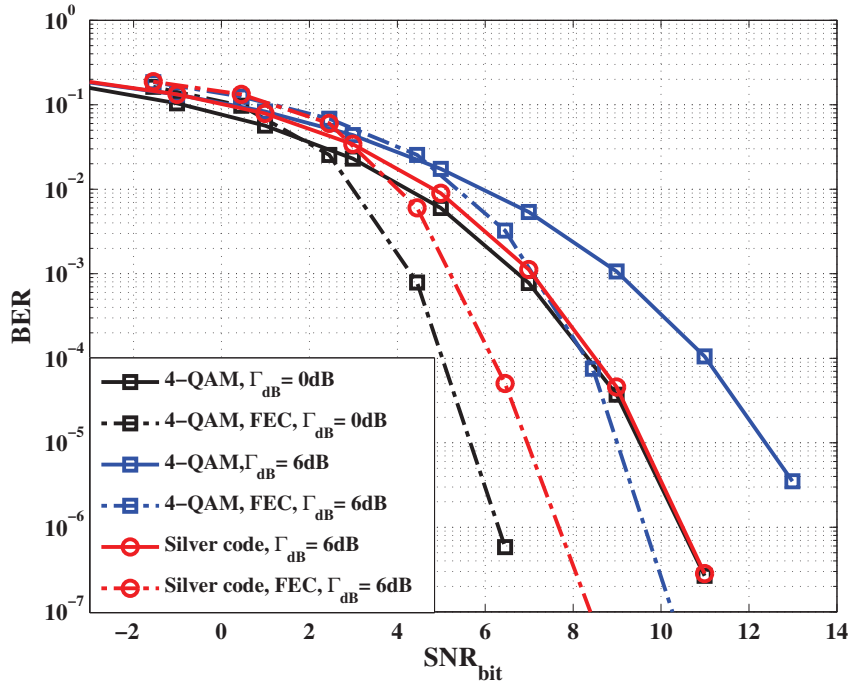
The bit error probability of a linear block code over an AWGN channel, and after hard decision decoding, is determined by the minimum distance of the considered FEC code $d_{min,FEC}$ and the crossover probability p of the equivalent binary symmetric channel (BSC) [11, Chap.7] as:

$$P_{e,HDD} \leq \sum_{m=t+1}^n \binom{n}{m} p^m (1-p)^{n-m} \quad (3.24)$$

where $t = \left\lfloor \frac{d_{min,FEC}-1}{2} \right\rfloor$ is the error correction capability of the linear block code and $d_{min,FEC}$ is the minimum Hamming distance between distinct codewords of this code. At high SNR, p tends



(a) HDD



(b) SDD

Figure 3.11: Bit Error Rate as a function of the SNR_{bit} for the uncoded scheme and the Silver code, with or without FEC: (a) Hard decision decoding (HDD), (b) Soft decision decoding (SDD). The simulated FEC is a BCH(63,45) code.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

towards zero and (3.24) is dominated by the first term where $m = t + 1$:

$$P_{e,HDD} \leq \binom{n}{t+1} p^{t+1} = \binom{n}{d_{FEC,HDD}} p^{d_{FEC,HDD}} \quad (3.25)$$

where $d_{FEC,HDD} = \left\lceil \frac{d_{min,FEC}}{2} \right\rceil$.

Had we considered soft decision decoding, the bit error probability would have been [11, Chap.7]:

$$P_{e,SDD} \leq (2^k - 1) p^{d_{FEC,SDD}} \quad (3.26)$$

where $d_{FEC,SDD} = d_{min,FEC}$.

p can be upper-bounded using (3.22) where the error probability is a codeword error probability. Given that a full-rate 2×2 PT codeword constructed with q -QAM symbols has $4 \log_2 q$ bits, an error in decoding one PT codeword implies that one bit is erroneous in the best case or all $4 \log_2 q$ bits are erroneous in the worst case, hence:

$$\frac{1}{4 \log_2 q} \Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \leq p \leq \Pr(\mathbf{X} \rightarrow \hat{\mathbf{X}}) \quad (3.27)$$

At high SNR, the PEP of the closest neighbors predominates the other terms in the union bound given in (2.27), and p is upper-bounded by:

$$p \leq A_{d_{PT}} \exp\left(-\frac{d_{PT}^2}{8N_0}\right) \quad (3.28)$$

where $d_{PT}^2 = \underset{\mathbf{X}, \mathbf{X}' \in \mathcal{C}}{\operatorname{argmin}} (\|\mathbf{X}_\Delta\|^2 - \gamma_{eq} \sqrt{a^2 + b^2})$ and $A_{d_{PT}}$ is the average number of codewords located at the distance d_{PT} of a given codeword. $A_{d_{PT}}$ is usually called the *kissing number*.

Replacing p by its upper bound and substituting the average energy per information bit E_b for the average symbol energy E_S , using $E_S = 1 = r_{PT} r_{FEC} E_b \log_2 q$, we get:

$$P_e \leq K(A_{d_{PT}})^{d_{FEC}} \exp\left(-\frac{d_{FEC} d_{PT}^2 \text{SNR}_{bit} r_{PT} r_{FEC} \log_2 q}{4}\right) \quad (3.29)$$

where K and d_{FEC} are the coefficient and the power of p in the bit error probability expression (3.24) or (3.26) depending on the chosen FEC decoding strategy. $r_{PT} = r_{ST}/n_t = r_{ST}/2$ is the symbol rate of the PT code per transmit polarization.

In order to evaluate the asymptotic gains provided by the concatenation of a FEC code and a PT code, we compare the following two schemes: a first scheme *NC* without FEC ($d_{FEC} = 1$ and $r_{FEC} = 1$) and without PT coding and a second scheme *FEC + PT* using a linear block code and a PT code. At the same achieved error probability, the effective coding gain G of the coded scheme

is given by:

$$G = \frac{\text{SNR}_{\text{bit}, \text{FEC}+\text{PT}}}{\text{SNR}_{\text{bit}, \text{NC}}} = \frac{r_{\text{FEC}} d_{\text{FEC}} r_{\text{PT}} d_{\text{PT}}^2}{d_{\text{NC}}^2} \quad (3.30)$$

In decibels, we obtain G_{dB} :

$$G_{\text{dB}} = 10 \log_{10}(r_{\text{FEC}} d_{\text{FEC}}) + 10 \log_{10} \left(\frac{r_{\text{PT}} d_{\text{PT}}^2}{d_{\text{NC}}^2} \right) = G_{\text{dB}, \text{FEC}} + G_{\text{dB}, \text{PT}} \quad (3.31)$$

The first term denotes the coding gain of the FEC code and the second term denotes the coding gain provided by PT coding. Equation (3.31) shows that the total asymptotic gain obtained when concatenating a FEC code and a PT code is the sum of the gains provided by each code separately, validating our numerical investigation.

Finally, we mention that apart from this theoretical analysis where we derived a closed form expression of the upper bound of the error probability explaining the performance of PT codes and the summation of their gains with the ones provided by FEC codes, we evaluated in a joint work with colleagues how far are the best coded schemes (PT + FEC) from the fundamental limit of PDL-impaired channels given by an outage probability analysis. We found that the use of the Silver code along with a powerful LDPC code construction provides a performance very close to the fundamental limit as long as PDL is less than 4 dB [175].

3.4 Experimental validation of PT coding gains

Until now, we investigated the mitigation of PDL using PT codes through numerical and theoretical analyses. This consideration was a necessary starting point due to the complexity of the optical channel and the difficulty to simultaneously model the interactions between all in-line effects. PDL being a distributed effect in the fiber, its statistics were studied in previous works and reviewed in the previous sections. The most commonly considered distribution of PDL is the Maxwellian resulting from the concatenation of components with low PDL with birefringent optical fibers. Moreover, other interactions can be taken into account such as interactions of PDL with distributed ASE noise [60] and non-linear effects [61].

The most appropriate way of considering these effects, all together, is to experimentally examine the ability of PT-codes to mitigate the PDL that results from distributed optical components in a long-haul link with distributed ASE noise. In a previous thesis, initial experimental results limited to the linear regime and to a lumped PDL element in a back-to-back configuration have shown the efficiency of PT codes [167, 168]. In this thesis, we aim to prove that a complete PT-coded OFDM system can be realized and is able to mitigate in-line PDL taking into account the various interactions with other channel effects.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

3.4.1 Experimental setup

To achieve this experimental demonstration, we built the following experimental setup in our laboratories and carried experimental measurements of BER curves and associated Q -factors. The optical link consists of a recirculating loop to emulate a 1000 km propagation. Fig. 3.12 shows the experimental setup.

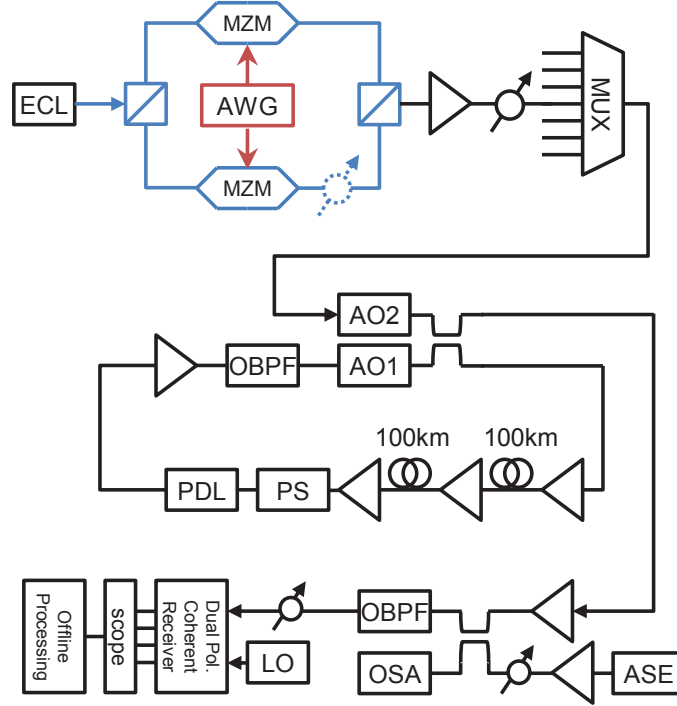


Figure 3.12: Experimental setup. (ECL: External Cavity Laser, AWG: Arbitrary Waveform Generator, MZM: Mach-Zehnder Modulator, MUX: Multiplexer, AO: Acousto-Optical Modulator, PS: Polarization Scrambler, OBPF: Optical Band-Pass Filter, ASE: Accumulated Spontaneous Emission source, LO: Local Oscillator, OSA: Optical Spectrum Analyzer, OSC: Tektronix 50GS/s Oscilloscope).

3.4.1.1 Transmitter side

The PDM-OFDM transmitter sends two OFDM signals on two orthogonal polarization states. Each OFDM symbol consists in 256 subcarriers including 194 data subcarriers and 10 pilot subcarriers for common phase noise estimation. A total effective bitrate of 25 Gb/s is achieved on a 9 GHz bandwidth determined by the DACs of the arbitrary waveform generator (AWG) providing the OFDM signals with a maximum sampling rate of 12 GS/s. The bitrate was fixed to 25 Gb/s in order to be in line with a 100 Gb/s four-band OFDM approach considered in [124]. The subcarriers are alternatively modulated with 4-QAM (No-coding), Silver-, Golden- and Alamouti-coded symbols. To ensure the same spectral efficiency for all codes, Alamouti-coded subcarriers use 16-QAM symbols whereas the Golden and Silver codewords are built using 4-QAM symbols.

An 18-sample cyclic prefix is appended to each OFDM symbol to absorb interference. A 22-symbols training sequence is inserted each 100 OFDM data symbols for time synchronization, frequency offset compensation and channel estimation. PT coding correlates the polarization-multiplexed symbols. Hence, four DACs are necessary to generate two complex-valued baseband OFDM signals. Having a unique AWG (Tektronix 7122B) with two independent outputs, we have to consider real-valued OFDM symbols satisfying the Hermitian symmetry property in order to implement a PT-coded scheme. This constraint does not affect our analysis knowing that the complex conjugate of the subcarriers will be treated independently of the original subcarriers at the receiver side. The raw bitrate is 36 Gb/s taking into account the different transmission overheads.

The up-conversion to the optical domain is done by two single-drive MZMs that we characterized and optimized their bias points (biased at the null point of their characteristic in order to ensure the best linearity in their transfer function). The optical carrier is generated by an external cavity laser (ECL) lasing at $\lambda = 1542.9$ nm and having a linewidth $\Delta\nu = 100$ kHz. Then, the two optical signals are combined by a polarization beam combiner and multiplexed among seven 50 GHz-spaced wavelengths that are not modulated. The PDM transmitter was also built by us and tested in a back-to-back configuration in order to accurately tune the path difference between the two tributaries of the PDM signal.

3.4.1.2 Optical link and receiver side

The transmission line consists in a recirculating loop which contains 2 spans of 100 km of SSMF, a home-built PDL element of ~ 2 dB, a polarization scrambler and an optical band-pass filter. The filter is needed to limit the noise accumulation in the loop and the polarization scrambler is necessary to fight important polarization dependent gains in the loop and generate performance results similar to the ones obtained on a deployed system. It also helps to emulate a time-varying channel. We used the polarization scrambler N7785B by Agilent Technologies at its maximum scrambling speed of 2474 rad/s which corresponds to a coherence time $t_c = 2\pi/2474 = 2.5$ ms way larger than the OFDM frame duration. One OFDM symbol duration is equal to the number of samples including the cyclic prefix multiplied by the sampling period: $(256 + 18)/(12.10^9) = 22.83$ ns, and an OFDM frame contains $100 + 22 = 122$ symbols which gives us a frame duration of $2.8 \mu s \ll t_c$.

The loop also contained variable optical attenuators after each EDFA in order to tune the gains. No more than 5 loops could be realized mainly because of the gain dependence on the number of channels (we had only 8 tunable lasers to create the WDM comb) and the absence of a differential gain equalizer in the setup. The loop was controlled by an active component comprised of two acousto-optic switches AO1 and AO2 and a 3 dB coupler. The AOs work in counter-phase as shown in Fig. 3.13. When AO1 is open and AO2 is closed, the signal is loaded in the loop. When

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

AO1 closes and AO2 is open, the signal circulates in the loop until AO1 opens again. The receiver is triggered to read data after 5 circulations. A pulse generator DG535 of Stanford Research Systems is used to control the AOs and trigger the receiver. The duty cycle of the pulse generator is adjusted to allow a complete filling of the loop in order to avoid gaps and hence strong transient effects in the EDFAs.

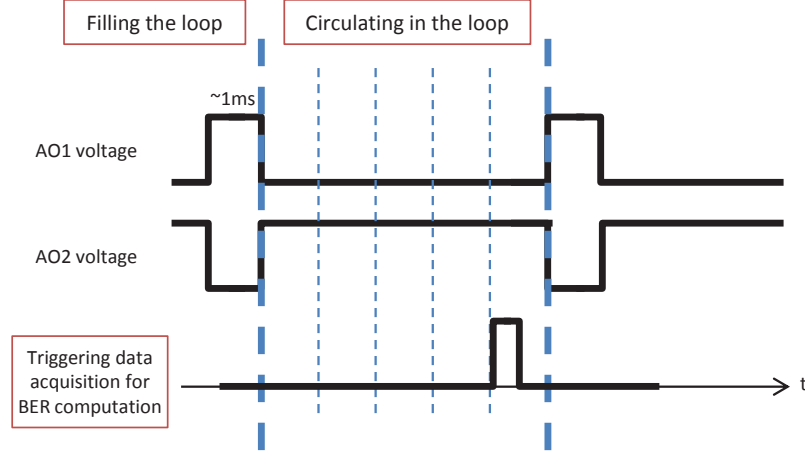


Figure 3.13: Acousto-optic switches for controlling the loop.

After additional ASE noise loading at the receiver, the desired wavelength is filtered and the signal is detected with a dual-polarization coherent receiver. The same ECL laser is used as an optical carrier at the transmitter and as a local oscillator at the receiver (intradyne detection). Two 90/10 couplers are also used at the receiver to monitor the signal on an optical spectrum analyzer and monitor the power inside the loop with a slow photodiode (photodiode + slow real-time oscilloscope). A real-time Tektronix DPO72004B oscilloscope is triggered to sample the received signal at 50 GS/s after five rounds in the loop.

$2 \cdot 10^6$ samples (10^5 bits/ PT code) are recorded for each channel realization and offline processing ending with ML decoding is carried to measure the BER and the corresponding Q -factors. Time synchronization is done using Schmidl and Cox algorithm, frequency offset correction (mainly to correct the drift of the ECL and the deterministic offset induced by the acousto-optic switches functioning at 40 MHz) is carried out using a frequency-domain method, phase error correction is made using the N_p pilot tones, and channel estimation is performed using an LSE estimator with the training sequence structure given in section 2.3.2.2. Finally, the estimated channel matrix is used to perform an exhaustive-search ML decoding followed by demodulation and a bit error count.

3.4.2 Experimental results

3.4.2.1 OSNR degradation with PDL

Before transmitting the PT-coded OFDM signals in the recirculating loop, we conduct preliminary measurements of the BER evolution as a function of the OSNR for 3 different PDL values at the transmitter: 0, 3 and 6 dB. An attenuator is inserted in one branch of the PDM transmitter to emulate the worst OSNR degradation of one polarization tributary. The loop is replaced by a polarization scrambler and the signal is then combined with an ASE noise source. The measured BER curves are plotted in Fig. 3.14. We firstly note that the Silver code performs better than the Golden code and secondly that the performance of the Alamouti code is independent of the PDL value as predicted by the numerical and theoretical studies.

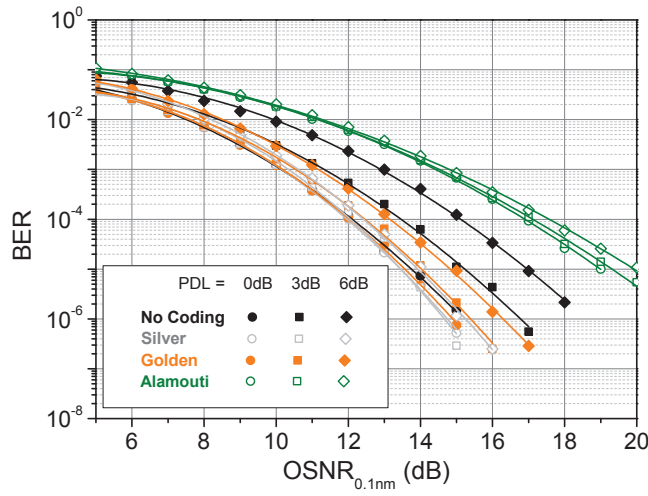


Figure 3.14: BER curves versus $\text{OSNR}_{0.1\text{nm}}$ with PDL = 0, 3, 6 dB for the different schemes: no coding, Silver, Golden and Alamouti codes.

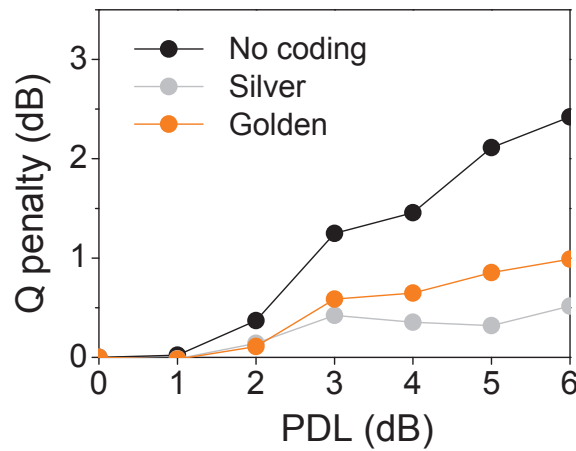


Figure 3.15: Q -penalty versus PDL at $\text{OSNR}_{0.1\text{nm}} = 13$ dB with regard to a PDL-free scenario.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS

In Fig. 3.15, for a PDL ranging from 0 to 6 dB, we show the Q -factor penalties obtained with 4-QAM, Silver and Golden codes with regard to a PDL-free case at an OSNR of 13 dB. We see that the Silver code performs the best reducing the Q -factor penalty by 2 dB for a PDL = 6 dB.

3.4.2.2 Lumped PDL

After these preliminary tests, we will transmit the PT-coded OFDM signals in the recirculating loop described in Fig. 3.12 and replace the in-line PDL element by an attenuator having the same insertion loss. PDL is emulated at the transmitter side as in the previous section. BER measurements are carried out after a 1000 km transmission for a span input power of the modulated channel ranging from -15 to 5 dBm. For each input power, up to 50 realizations are considered depending on the target BER. The obtained bathtub-shaped curves are plotted in Fig. 3.16 for the 4-QAM and Silver-coded schemes at a PDL of 0, 3 and 6 dB. We firstly note that all curves have their optimum operating point at -3 dBm. However, the uncoded scheme severely suffers from PDL ($Q_{\text{dB}} = 10.5$ dB at -3 dBm for PDL= 6 dB) while the Silver code completely compensates the penalty at PDL=3 dB and nearly all penalty at PDL=6 dB ($Q_{\text{dB}} = 12.3$ dB at -3 dBm). Moreover, an important observation is that the PT-coded OFDM does not induce any extra penalties in the non-linear propagation regime as predicted by the numerical simulation.

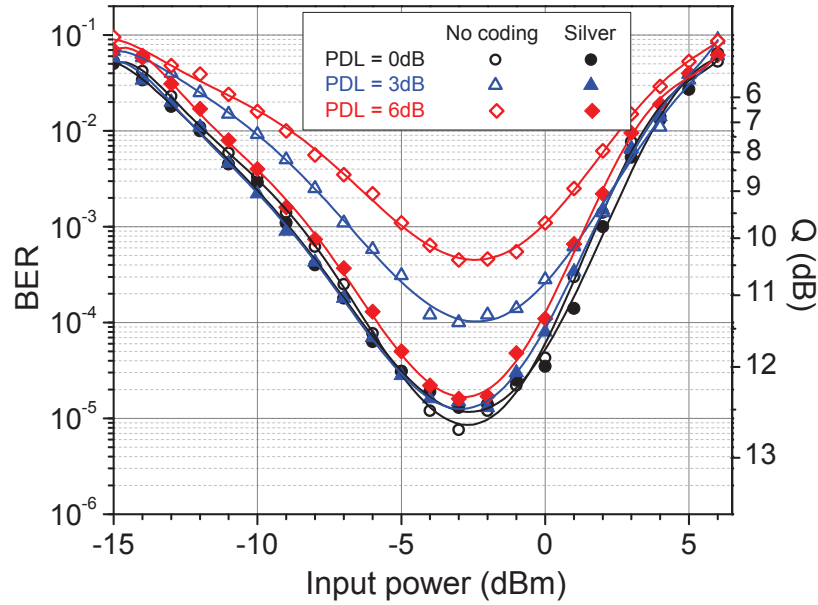


Figure 3.16: Comparison of BER evolution versus launched input power after 1000 km for the Silver-coded and QPSK schemes (PDL at the transmitter = 0, 3, 6 dB).

3.4.2.3 Distributed in-line PDL

PDL being a distributed effect induced by in-line components along the optical transmission link, we are interested in testing the capability of codes in mitigating in-line PDL. Therefore, we remove the emulated PDL at the transmitter and insert a PDL element of ~ 2 dB into the loop. Then, we extract the Q -factor distribution for the different coding schemes by measuring 2000 Q -factor values at the optimum operating point (NLT) corresponding to an input power of -3 dBm. The measured OSNR at the receiver is 12 dB after additional ASE noise loading in order to evaluate a complete range of Q -factors.

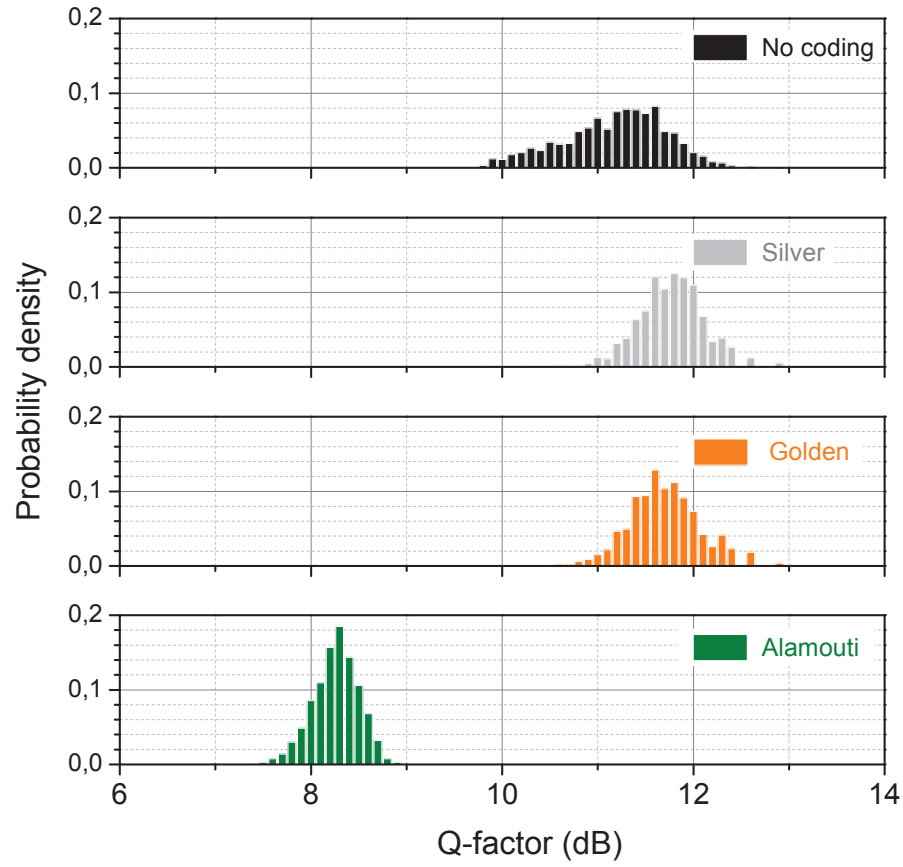
The obtained Q -factor distributions are shown in Fig. 3.17a. After 5 loops, we measure PDL using the estimated 2×2 channel matrices and obtain a Maxwellian probability distribution function of PDL with a 4.2 dB mean averaged over all the subcarriers (the expected theoretical mean being approximately $2\text{dB}\sqrt{5} = 4.5$ dB [53]). Fig. 3.17b shows the measured probability distribution and the fitted Maxwellian distribution.

We notice that the Q -factor distribution for the 4-QAM-coded subcarriers has a mean of 11.2 dB and a standard deviation (std) of 0.55 dB whereas the mean Q -factor of the Silver- and Golden-coded subcarriers is 11.8 dB (the same observed value for the PDL-free case when OSNR = 12 dB). The Q -factor distributions are also narrower with PT coding. Note, as well, that the Silver code gives a distribution slightly narrower than the one obtained with the Golden code: std of 0.35 dB and 0.37 dB respectively.

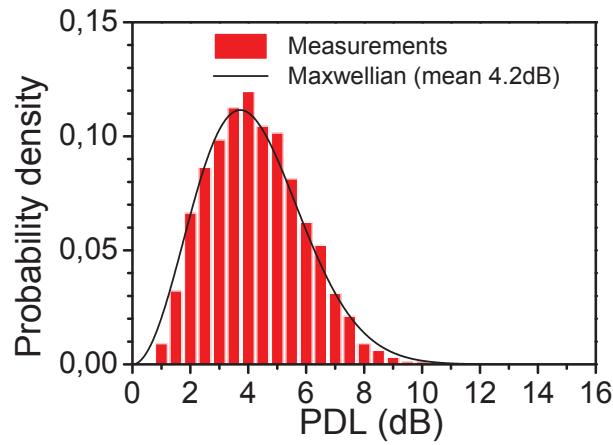
Regarding the Alamouti code, we observe a mean Q -factor of 8.3 dB due to the use of 16-QAM symbols instead of 4-QAM to guarantee the same spectral efficiency. However, its Q -factor distribution is the narrowest of all with a std of 0.24 dB. This observation can be explained by the powerful orthogonal structure of the Alamouti codeword matrix that makes its performance independent of the amount of accumulated PDL in the link. The observed small variance of its Q -factor distribution can result from the increased sensitivity of 16-QAM modulation to phase noise and non-linear effects.

In conclusion, we have experimentally demonstrated, for the first time, the potential of PT codes to mitigate PDL in long-haul optical links, as well as their resistance to non-linear effects. An enhanced performance evaluated through Q -factor measurements is observed when PT codes are used: a narrower Q -factor distribution is obtained and the mean Q -factor is improved.

3. POLARIZATION-TIME CODING FOR PDL MITIGATION IN POLARIZATION MULTIPLEXED SYSTEMS



(a) Q -factor probability distributions after 1000km for $P_{in} = -3\text{dBm}$ ($OSNR_{0.1nm} = 12\text{dB}$).



(b) Probability distribution of PDL measured with 2000 realizations.

Figure 3.17: Experimental results.

Summary

We have presented numerical and theoretical studies of PDL mitigation using redundancy-free coding schemes. The Silver code can efficiently mitigate the PDL in current long-haul optical links where PDL values rarely exceed 6 dB. However, the theoretical study of PDL mitigation showed that this code is not optimal. The established upper bound of the error probability for an optical channel with PDL yields the design criterion required to construct codes that guarantee a PDL-independent error probability, restoring the performance over an additive white Gaussian channel. These design rules open the way for the construction of optimal codes dedicated to PDL mitigation in the optical channel. Besides, we have also shown that the total coding gain provided by the concatenation of a FEC code and a PT code is the sum of the separate coding gains. A transmission experiment with distributed in-line PDL was accomplished as well, validating the results obtained with a simplified channel model. The investigated PT codes do not induce extra penalties when non-linear effects are considered and they can improve both the mean and the variance of the Q -factor distributions. In the next chapter, we will move to larger mode-multiplexed optical MIMO systems to explore the benefits of ST coding in mitigating penalties induced by non-unitary effects known as mode dependent loss (MDL).

Chapter 4

ST coding for MDL Mitigation in Mode Division Multiplexed (MDM) Systems

The main goal of the study of PDL mitigation using ST codes was understanding the behavior of MIMO coding techniques on a 2×2 optical channel impaired by non-unitary effects. The fading nature of this channel differs from the one of the wireless channel. Simulations and experimental results revealed important differences in the performance of ST codes. Our theoretical analysis proved these observations showing that ST codes on an optical transmission channel are not meant to bring any diversity gains. They only enhance the minimum distance of two different codewords decreased by PDL. This is mainly due to the fact that both polarizations cannot be in fading at the same time because the total signal power, on both polarizations, is preserved by periodical optical amplification. These findings were an essential starting point for understanding non-unitary effects in larger optical MIMO systems, specifically mode division multiplexed systems.

MDL is a non-unitary effect impairing SDM systems in the same way as PDL affects PDM systems. However, MDL scenarios are much more diverse because different combinations of gain disparities may exist between spatial modes $M > 2$ depending on the sources of MDL in the optical link while PDL is defined as a single gain disparity between two polarization modes. For instance, two different technologies of few-mode amplifiers (FMAs) may show different mode dependent gain (MDG) values by privileging different spatial modes. Moreover, different SDM coupling strategies are envisaged in current research works: on the one hand, a low-modal coupling strategy is suggested to avoid full MIMO processing of the spatial modes at the receiver, hence applying smaller 2×2 MIMO filters to separate the polarization modes of each spatial mode that does not exhibit a degeneracy (i.e. LP_{0v} modes) or else 4×4 MIMO filters to separate the polarization modes and the degenerate modes (i.e. $LP_{\mu va}$ and $LP_{\mu vb}$ modes with $\mu \geq 1$). On the other hand, a strong-coupling strategy is also possible requiring a $2M \times 2M$ MIMO processing. While having a more complex digital receiver, the latter strategy benefits of strong coupling to

reduce the maximum differential delays and losses between the modes and will be adopted in this chapter.

In this chapter, we will focus on the management of MDL in FMF-based coherent optical transmission systems. This focus is driven, on one hand, by the emergence of new fabrication processes of few-mode optical components such as few-mode fibers with low DMGD [176] and few-mode amplifiers (FMAs) with an optimized mode dependent gain (MDG) [92]. Conventional MMFs with large core diameters support an important number of modes with large propagation delays which makes them unsuitable, for the moment, for metro and long-haul transmissions leading most research works to consider FMFs with low DMGD supporting 3 or 6 modes [82, 83]. Once the design of MMFs and single-mode or multi-mode MCFs gains more maturity and becomes practical, we will have to deal with similar non-unitary crosstalk issues arising from the fiber and the integrated optical devices.

Before simulating ST-coded transmissions over 3-mode and 6-mode systems, we start by briefly reviewing the description of spatial modes in graded-index few-mode fibers with a parabolic profile in section 4.1, then we look into the MDL-induced penalties in three different SDM channel models in section 4.2. First, we recall a simulation model suggested by Kahn et al. [74] depicting MDL statistics in a strong coupling scenario. Later, we focus on two more general models where two different MDL sources are studied: misalignments at splices and connections that induce non-unitary crosstalk and MDG at few-mode amplifiers. For both, we evaluate the statistics of MDL and list the optical solutions that were suggested to reduce MDL. Afterwards, in section 4.3, we evaluate the potential of ST coding in absorbing MDL-induced penalties in the various channel models. Finally, in section 4.4, we analyze the complexity of the proposed ST coding solution to get a sense of their scale-up possibilities to larger SDM systems.

4.1 Description of LP modes in parabolic-index FMFs

All over the chapter, we will consider graded-index fibers with a concave parabolic index profile (red curve in Fig. 4.1), a core radius r_c and a numerical aperture $NA = \sqrt{n_{co}^2 - n_{cl}^2} = 0.205$ where n_{co} is the maximum refractive index in the core and n_{cl} the refractive index in the cladding of the fiber, and we operate at $\lambda = 1550$ nm. The choice of graded-index fibers with a concave parabolic profile is based on the small differential modal group velocities that can be achieved with these fibers compared to the ones obtained with step-index profiles, and hence the minimal DGD spread between the modes as noted in section 1.4.4.

The number of propagating spatial modes in a fiber depends on its V parameter, a dimensionless variable determined by opto-geometrical properties of the fiber: the numerical aperture NA , the wavelength λ and the core radius r_c . The LP_{01} mode is the fundamental mode with a null cut-off frequency. For a given λ and NA , when increasing r_c , the cut-off frequency of the LP_{11} mode

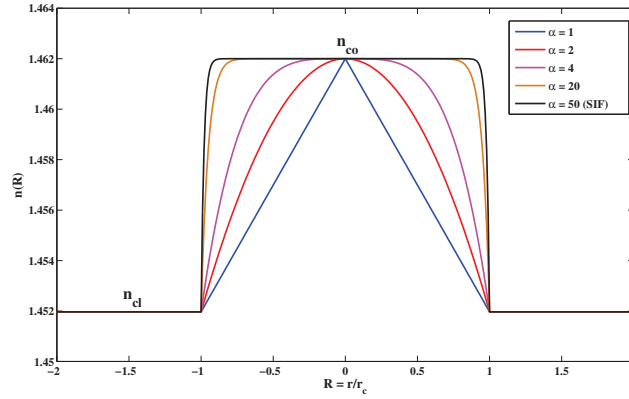


Figure 4.1: Various graded-index profiles with a refractive index $n(R) = n_{co} (1 - 2\Delta R^2)^{1/\alpha}$ where $\Delta = (n_{co} - n_{cl})/n_{co} \ll 1$ and $R = r/r_c$ the normalized radius. Special cases: $\alpha = 2$ parabolic index, $\alpha = \infty$ step-index.

$V_c = 3.52$ is reached [177]. This mode exhibits a degeneracy into two orthogonal modes: LP_{11a} and LP_{11b} resulting in a total of three spatial modes. Increasing the radius r_c more, the cut-off frequency of three modes will be reached: $V_c = 5.1$ for the LP_{02} mode and $V_c = 5.7$ for the two degeneracies of the LP_{21} mode [177]. For $NA = 0.205$ and $\lambda = 1550$ nm, a core radius of $6 \mu\text{m}$ (resp. $8.7 \mu\text{m}$) can guarantee a 3-mode (resp. 6-mode) operation.

In a weakly-guiding graded-index fiber with an infinite parabolic profile index ($\alpha = 2$) such that $\Delta = (n_{co} - n_{cl})/n_{co} \ll 1$, the field distributions of the linearly polarized (LP) modes can be approximated by Laguerre-Gauss polynomials and are given by [18, Chap.5]:

$$E_{\mu,v}(r, \phi) = C_{\mu,v} \left(\frac{r}{w} \right) L_v^{(\mu)} \left(\frac{r^2}{w^2} \right) \exp \left(\frac{-r^2}{2w^2} \right) \begin{Bmatrix} \sin(\mu\phi) \\ \cos(\mu\phi) \end{Bmatrix} \quad (4.1)$$

where r and ϕ are the polar coordinates, μ is the circumferential order (non-negative integer) and v the radial order (strictly positive integer) of $LP_{\mu v}$, $C_{\mu,v}$ a normalization factor chosen to fulfill $\iint_A E_i \cdot E_j dA = \delta_{i,j}$, $L_v^{(\mu)}(x)$ the generalized Laguerre polynomials of order μ and degree v and $w = \sqrt{r_c/k_0 NA}$ the spot size of the fundamental mode and $k_0 = 2\pi/\lambda$ the free space wavenumber. The curly brackets denote the two degeneracies of the modes with $\mu \geq 1$. While in a realistic fiber the profile is a truncated parabola, the LP modes are still well approximated by (4.1). The spatial distributions of the first 6 modes are shown in Fig. 4.2.

4.2 MDL in MDM transmission systems

Based on previous works [2, 74, 78, 90], we define channel models to simulate various scenarios of MDL-impaired mode-multiplexed transmissions. We observe the statistics of MDL and assess the performance of each system. In order to solely focus on modal loss disparities, single-polarization

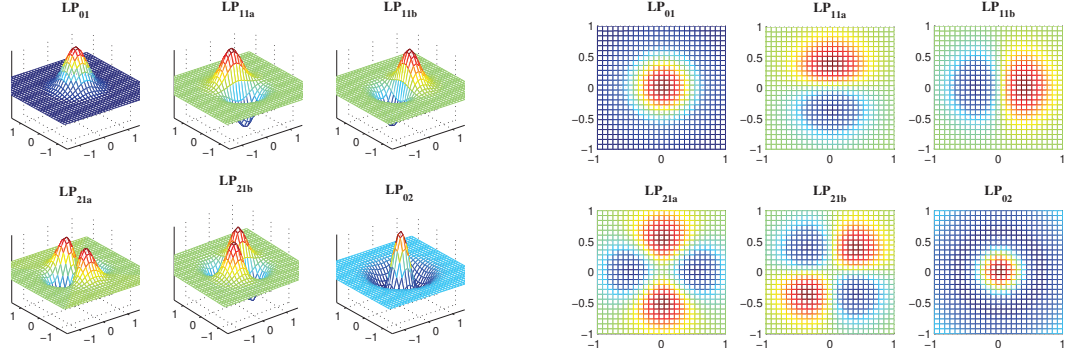


Figure 4.2: Spatial distributions of the LP_{01} , LP_{11a} and LP_{11b} , LP_{21a} and LP_{21b} , and LP_{02} modes given by the Laguerre-Gauss approximation: 3D view on the left, and 2D view of (x, y) plane on the right.

SDM schemes are considered and non-linear effects are neglected as in the case of many previous studies on MDL [78, 90]. We also neglect any frequency dependence of MDL and any coloration of noise in the following models.

Had the crosstalk in the MDM channel been unitary, meaning that the paths remain orthogonal after propagating through the optical channel, separating them and removing ISI induced by the differential delays, would have been enough in order to retrieve the performance of a crosstalk-free channel. However, long-haul MDM systems are impaired by modal loss disparities, also known as MDL, arising from imperfections in the link such as non-unitary couplings in the fiber due to bending losses and splices [90] as well as mode dependent gains at optical amplifiers [68, 92]. MDL deteriorates the overall system performance [2, 90] because the modes are no more orthogonal and experience different levels of attenuation while propagating. Unlike dispersion and unitary crosstalk, MDL is a capacity-limiting effect that can reduce the number of propagating modes available for multiplexing [64, 74]. Hence, it is definitely a challenge that needs to be surmounted.

4.2.1 Best-case scenario: strong-coupling model

The first model, proposed and studied by Ho and Kahn [74], illustrates the accumulation of MDL under the hypothesis of strong coupling in the optical fiber and components. The optical link consists of a concatenation of $K > 20$ independent sections with MDL, given by:

$$\begin{aligned} \mathbf{H}_{M \times M} &= \prod_{k=1}^K \mathbf{M}_k = \prod_{k=1}^K \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{V}_k \\ &= \mathbf{U} \text{diag}[g_1, g_2, \dots, g_M] \mathbf{V} \end{aligned} \quad (4.2)$$

where $\mathbf{U}_k, \mathbf{V}_k$ are random unitary matrices and $\lambda_{i=1 \dots M}^{(k)} = \exp(\alpha_i^{(k)}/2)$ are the diagonal elements of $\mathbf{\Lambda}_k$. Two metrics are defined to evaluate MDL: the standard deviation (std) of all modal gains

$g_i, i=1\dots M$ in dB $\sigma_{MDL} = \sqrt{\mathbb{E}[g_i^2]}$ and the maximum gain difference $MDL = \max(g_i)/\min(g_i)$ in dB. By gain, we mean the coefficients in the diagonal matrix obtained by an SVD of $\mathbf{H}$ that can be smaller or larger than 1.

The simulation of this model in [74] is done only for an even number of modes M and with half the modes having a gain α and the other half a gain of $-\alpha$. The simulation results show that for low to moderate MDL levels, the distribution of the modal gains in dB is identical to the distribution of the eigenvalues of a specific random matrix family, the zero-trace Gaussian unitary ensemble (GUE) [74], and the standard deviation of modal gains is given by:

$$\sigma_{MDL} = \frac{10}{\log 10} \xi \sqrt{1 + \xi^2/12} \quad (4.3)$$

where $\xi = \sqrt{K}\alpha$ is the square root of the accumulated MDL variance in unit of log power gain. For $M = 2$ modes, the obtained distribution of the modal gains is Maxwellian as the distribution of PDL in PDM systems. For a higher number of modes, the obtained distribution has as many peaks as the number of modes and for an extremely high number of modes, the distribution converges to a semi-circle distribution [74]. We checked these findings for $M = \{2, 4, 8, 64\}$ and compared the modal gains obtained from (4.2) with the eigenvalues of the equivalent zero-trace GUE generated as explained in [74]. The results are shown in Fig. 4.3 for $\xi = 1$ or equivalently $\xi_{dB} = 10\xi/\log 10 = 4.34$ dB. We found that the approximation is accurate for a large number of modes when $K > 32$ and for std of overall modal gains up to 15 dB as noted in [74].

We also show in Fig. 4.4 the obtained distribution of the maximum MDL difference defined as the ratio between the highest and lowest gains for the different numbers of modes. For the same variance of the accumulated MDL $\xi^2 = 1$, the average maximum MDL grows with the number of modes and approaches a Gaussian distribution for $M > 2$ which was assumed by Winzer *et al.* in [64].

The most important conclusion to retain from these results is the reduced MDL accumulation in a strong coupling scenario where the std of accumulated modal gains scales as a function of the square root of the number of MDL sources K in the link instead of linearly scaling with K in a weak or no coupling scenario, resulting in an ill-conditioned channel matrix (i.e. a matrix with a very large ratio between the highest and lowest singular value). The authors in [74, 84] underline the benefits of strong coupling for the reduction of both DMGD and MDL spread, hence the reduction of the equalization complexity of ISI and the penalties of MDL. In the same works, the authors point out that, in presence of DMGD, the accumulated MDL can be frequency dependent depending on whether the signal bandwidth is smaller or larger than the coherence bandwidth of the channel. However, as in the case of PDL, this frequency dependence will reduce the variance

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

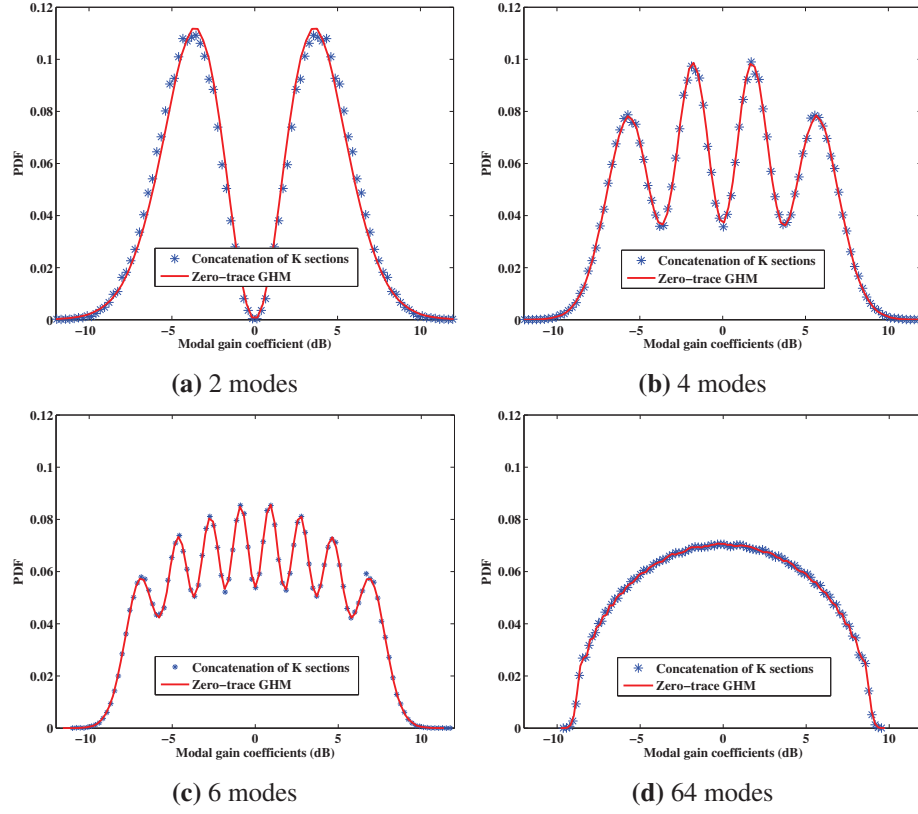


Figure 4.3: Probability distribution functions of the modal gains for different number of sections $K = \{32, 64, 128, 256\}$ with $\xi = 1$ in units of logarithm of power gain compared to the eigenvalues of the equivalent zero-trace GUE.

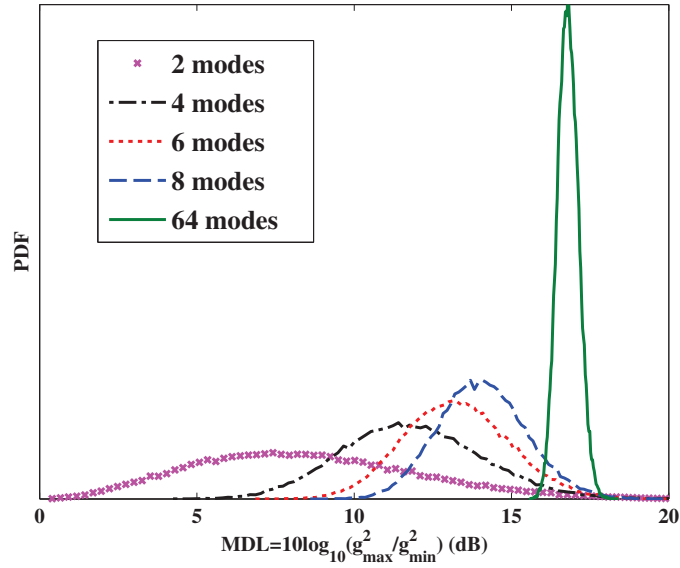


Figure 4.4: Probability distribution functions of the maximum MDL for $M = \{2, 4, 6, 8, 64\}$.

of the instantaneous capacity and BER of the link, hence we will maintain a frequency independent model of MDL.

While this model provides an elegant statistical representation of MDL through the equivalence to the eigenvalue distribution of a zero-trace GUE, it has some limitations starting with the coupling scenario where only strong coupling is depicted and also with the constraint of an even number of modes. Another close model is proposed by Winzer *et al.* in [64] where the individual gains in the sections are drawn from a uniform distribution. The authors also note that the maximum accumulated MDL scales with the square root of the K MDL elements in the link in the strong coupling regime. However, both models are limited to the strong coupling case and do not take into account an underlying fiber design or real channel imperfections. Hence, the conclusions drawn from these models may differ from realistic SDM channels. In the following, we study two SDM channel models that take into consideration fiber and optical component imperfections that we may observe in an SDM optical link.

4.2.2 Case study 1: fiber misalignments and micro-bends

4.2.2.1 Channel model

In [90], an MMF-based SDM system with M propagating modes is proposed where discrete linear modal crosstalk is introduced at imperfect connectors and splices as shown in Fig. 4.5. In the same work, the periodic insertion of mode scramblers was also proposed to average the losses experienced by the modes as shown in Fig. 4.6. Polarization crosstalk was not addressed. Neglecting any fiber non-linearity, the resulting MIMO channel model is given by:

$$\mathbf{Y}_{M \times T} = \mathbf{H}_{M \times M} \mathbf{X}_{M \times T} + \mathbf{N}_{M \times T} = \sqrt{\alpha} \prod_{k=1}^K (\mathbf{T}_k \mathbf{C}_k \mathbf{P}_k) \mathbf{X}_{M \times T} + \mathbf{N}_{M \times T} \quad (4.4)$$

where $\mathbf{X}_{M \times T}$ (resp. $\mathbf{Y}_{M \times T}$) are the emitted (resp. the received) symbols on the M modes and during T time slots, $\mathbf{H}_{M \times M}$ is the $M \times M$ channel matrix that consists of a concatenation of K fiber sections. The fibers are modeled with a diagonal $M \times M$ matrix $\mathbf{T}_k$ with random phase entries $\exp(i\phi_m)$ and $\phi_m \in [0 : 2\pi]$. DMD is not considered since it does not affect the capacity of the system and can be equalized using time domain filters or an OFDM format with a suitable cyclic prefix [104].

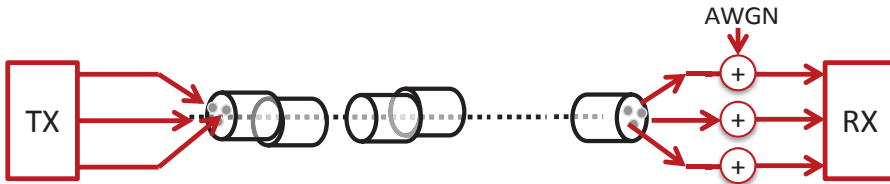


Figure 4.5: SDM channel model impaired by fiber misalignments and micro-bends.

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

$\mathbf{C}_k$ represents a random non-unitary modal coupling due to a fiber misalignment at a splice or a connector. The elements $c_{i,j}$ of the $M \times M$ matrices are computed using an overlap integral of the electrical field distributions of the propagating modes over the fiber cross section as in [90]:

$$c_{i,j} = \iint_A E_i(x,y) E_j^*(x+\Delta x, y+\Delta y) dA \quad (4.5)$$

where $\{i, j\} = \{1 \dots M\}$, A is the fiber cross section, E_i the normalized complex field of the mode i and $\Delta x, \Delta y$ the fiber misalignment assumed to be independent Gaussian distributed in directions x and y with zero mean and standard deviation $\sigma_{x,y}$. In [78], the authors also used misalignment matrices to model the bending losses in fibers. They showed that a misalignment matrix can also depict the energy loss in higher order non-propagating modes when the fiber has bends. We create a database of the matrices $\mathbf{C}_k$ computed over a grid of misalignments in the x and y directions with a step of $0.2\%r_c$ and using the spatial distributions of the modes given in (4.1).

$\mathbf{P}_k$ correspond to mode scramblers and are replaced by a random $M \times M$ permutation matrix when k is a multiple of a certain scrambling period denoted by K_{scr} , or an identity matrix otherwise. α is a normalization factor used to compensate the common link loss. Finally, $\mathbf{N}_{M \times T}$ is the noise assumed to be additive white Gaussian with variance N_0 per real dimension.

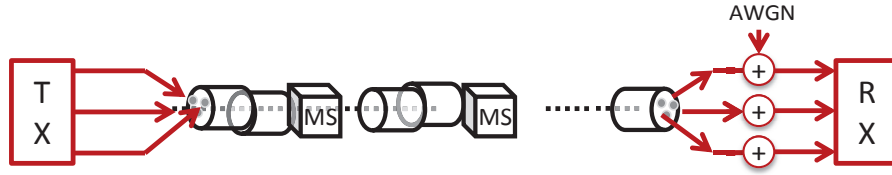


Figure 4.6: SDM channel model impaired by fiber misalignments and micro-bends with periodic mode scrambling (MS).

In a conventional MMF-based OFDM-SDM transmission system shown in Fig. 4.7, spatial multiplexing (SM) is performed by sending independent data streams on each mode at each time slot, hence $T = 1$ and the codeword $\mathbf{X}_{M \times 1}$ is a vector of independent q-QAM symbols. In presence of non-unitary impairments such as MDL, a simple ZF equalizer will result in a sub-optimal performance due to the noise enhancement by the channel inversion performed at the receiver. In order to optimally decode the received information, the ML criterion should be satisfied. In the 2×2 PDM schemes, the ML criterion was implemented through exhaustive search among all possible codewords. However, the complexity of this technique increases exponentially with the size of the MIMO system as noted in chapter 2. Hence, a sphere decoder [156], a reduced search technique, will be used for SDM systems. The corresponding algorithm can be found in appendix D. The initial radius of the sphere decoder is initialized as a function of the noise variance and the

lowest eigenvalue of the channel as explained in section 2.4.4. The sphere is centered at the received point and the algorithm searches, in the sphere, for the closest symbol to the received point starting from the surface of the sphere and going towards its center. If no symbols are found in the initial sphere, the radius is increased and the search is carried out in a new larger sphere in order to satisfy the ML criterion.

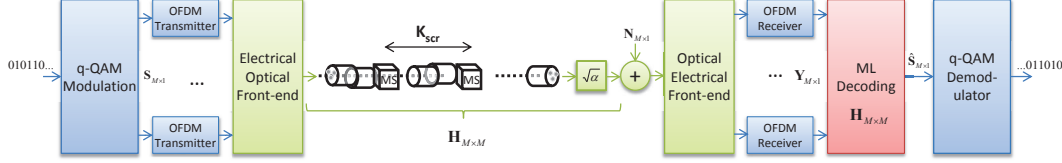


Figure 4.7: Mode-multiplexed OFDM optical transmission system.

4.2.2.2 MDL-induced penalties

We consider an SDM system as defined in (4.4) with $M = 3$. At the transmitter, we send 4-QAM symbols $s_{m=1,2,3}$ of unit energy $E_S = 1$ over each mode, providing a spectral efficiency of six bits per channel use. We suppose that graded-index fibers with a parabolic index profile, a core radius $r_c = 6 \mu\text{m}$ and a numerical aperture $NA = 0.205$ are used. With these parameters and operating at a wavelength $\lambda = 1550 \text{ nm}$, three LP modes are allowed to propagate corresponding to the modes in the first row of Fig. 4.2.

	LP_{01}	LP_{11a}	LP_{11b}
LP_{01}	-0.01	$-\infty$	-26.50
LP_{11a}	$-\infty$	-0.01	$-\infty$
LP_{11b}	-26.50	$-\infty$	-0.03

(a) $\Delta y = 3\%r_c$

	LP_{01}	LP_{11a}	LP_{11b}
LP_{01}	-0.01	-26.50	$-\infty$
LP_{11a}	-26.50	-0.03	$-\infty$
LP_{11b}	$-\infty$	$-\infty$	-0.01

(b) $\Delta x = 3\%r_c$

	LP_{01}	LP_{11a}	LP_{11b}
LP_{01}	-0.01	-30.02	-30.02
LP_{11a}	-30.02	-0.02	-60.03
LP_{11b}	-30.02	-60.03	-0.02

(c) $\Delta x = 2.1\%r_c, \Delta y = 2.1\%r_c$

Table 4.1: Loss and crosstalk levels in dB between the LP_{01} , LP_{11a} and LP_{11b} modes at the output of three different misalignment scenarios.

The crosstalk coefficients $c_{i,j}$ at the misalignments are computed using the field distributions of the LP modes. We infer from the different symmetries in these distributions shown in Fig. 4.2 that each mode is more or less impacted than the others by a misalignment of a given amplitude and orientation. In table 4.1, we show the loss and crosstalk levels in dB of the coefficients of three

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

specific misalignment matrices corresponding respectively to a vertical misalignment $\Delta y = 3\%r_c$, a horizontal misalignment $\Delta x = 3\%r_c$, and a $3\%r_c$ misalignment at 45 degrees.

The simulated channel consists of $K = 300$ misaligned fiber slices with a random Gaussian misalignment of zero mean and standard deviation $\sigma_{x,y} = \{2\%, 3\%, 4\%\}r_c$ in the x and y directions. The scramblers were inserted in the channel with scrambling periods set to $K_{scr} = \{1, 16\}$. In Fig. 4.8, we show the statistics of the MDL, defined as the ratio between the maximum and the minimum singular values of the channel, that result from the concatenation of 300 misaligned sections, for various misalignment levels and both with or without mode scramblers. 10^6 channel realizations were drawn for each scenario. Had we considered an std lower than $2\%r_c$, the obtained channel would have been almost unitary and for an std greater than $4\%r_c$, the accumulated MDL would have become very large.

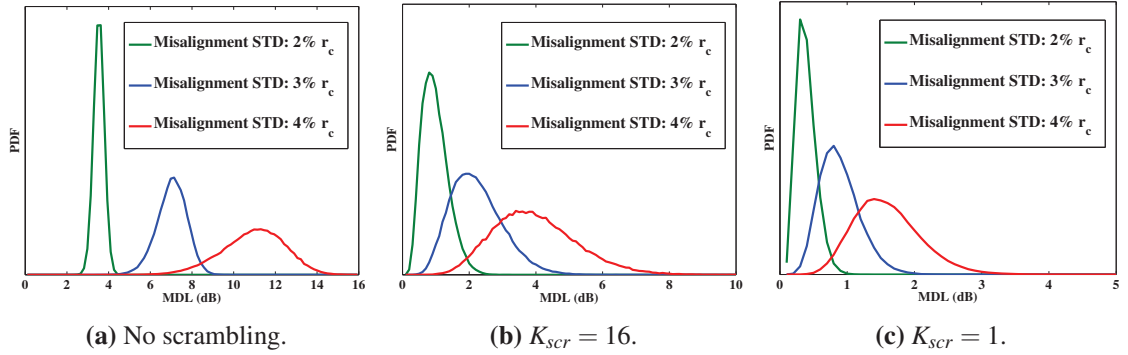
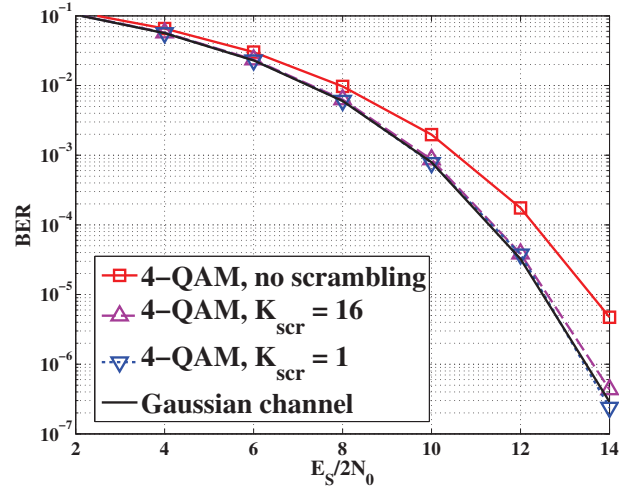
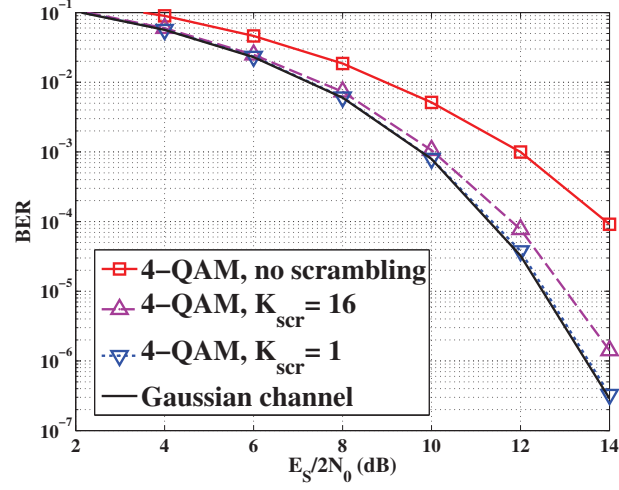
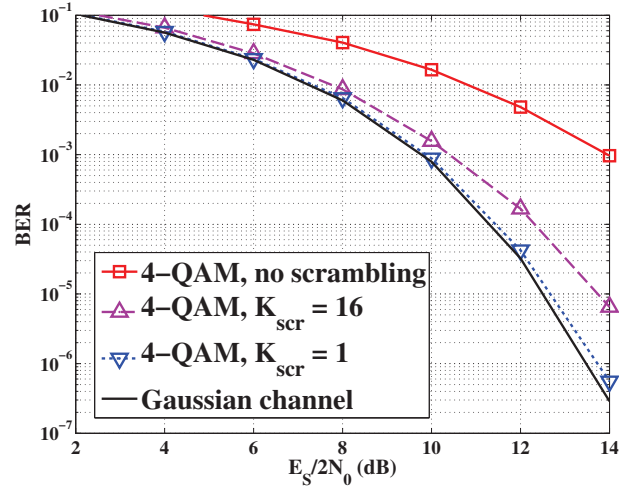


Figure 4.8: PDF of MDL defined as the ratio between the highest and the lowest singular values of the 3×3 channel for different misalignment levels and various mode scrambling scenarios ($K = 300$).

Then, we measure the performance, in terms of BER curves versus the $\text{SNR} = E_s/2N_0$, of the spatial multiplexing scheme with and without the scrambling option. The obtained performance of the ML sphere decoder is shown in Fig. 4.9. From the square marked red curves in Fig. 4.9a, 4.9b and 4.9c, we notice that the SNR penalty induced by fiber misalignment increases from 0.8 dB for a misalignment standard deviation (std) of $2\%r_c$ to 4 dB for a misalignment std of $4\%r_c$, at a $\text{BER} = 10^{-3}$. For a misalignment std lower than $2\%r_c$, the channel is almost unitary and no penalty is induced whereas for an std larger than $4\%r_c$, the penalties become extremely severe. Indeed, at a higher misalignment, the resulting accumulated MDL is also higher and the performance degradation increases as can be seen from the obtained MDL statistics shown in Fig. 4.8.

When mode scramblers are inserted each 16 fiber sections, the SNR penalty at $\text{BER} = 10^{-3}$ is fully absorbed at $2\%r_c$, a small penalty of 0.3 dB remains at $3\%r_c$ and 0.7 dB at $4\%r_c$. When scrambling at each section, the penalty is fully absorbed in all cases. Scrambling reduces MDL by averaging the losses over all the modes as can be seen in Fig. 4.8 from the statistics of MDL and as demonstrated in [90, 94]. Note that when the modes are fully coupled by periodical scrambling, MDL is reduced but not eliminated.

(a) $\sigma_{x,y} = 2\%r_c, SM$ (b) $\sigma_{x,y} = 3\%r_c, SM$ (c) $\sigma_{x,y} = 4\%r_c, SM$ **Figure 4.9:** BER curves of the ML-decoded 3×3 MDM system (spectral efficiency = 6 bits/cu).

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

When scramblers are installed at each misalignment, a strong coupling regime is reached and the average MDL grows with the square root of the number of MDL sources [64, 90]. In strong coupling, the modal gains will also follow the statistics of the first model proposed by Kahn *et al.* and discussed in the previous section. We looked into the statistics of the modal gains in order to verify that their variance is minimized and their PDF tends to a distribution with three peaks, when intensive mode scrambling is used. This is verified and illustrated in Fig. 4.10 for a misalignment of $4\%r_c$. According to [90], this misalignment value induces a std of modal gain coefficients of 0.05 dB per misalignment, which corresponds to a square-root of the accumulated MDL variance $\xi_{dB} = \sqrt{K} \times 0.05 = \sqrt{300} \times 0.05 = 0.866$ dB or $\xi = 0.1994$ in unit of log power in the strong-coupling model.

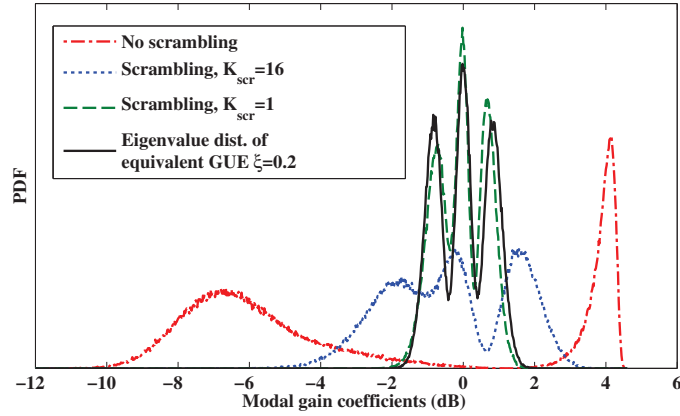


Figure 4.10: PDFs of modal gains of the 3×3 channel for a misalignment of $4\%r_c$ and various mode scrambling scenarios.

In [90], the authors simulate the overall MDL for various numbers of modes as a function of the average splice loss generated by a misalignment. They point out that splice losses as low as 0.03 dB may lead to an outage after few hundred kilometers, considering that a splice is present each 1 or 2 km. They also observe through the simulation results that, for different numbers of modes, the capacity reduction solely depends on the total accumulated splice loss that should be less than ~ 2 dB to limit the capacity reduction to 90% of the capacity of the MDL-free channel. There are currently no misalignment limits defined in the splicing norms for FMFs. In manuals of Corning fibers, 0.05 to 0.1 dB loss per slice is tolerated for SMFs. Recently, the feasibility of handling splices for FMFs with conventional techniques was demonstrated in [178]. A reproducible low splice loss of 0.01 to 0.05 dB and low coupling losses were observed for 3- and 6-mode FMFs. While such improvements may reduce the induction and accumulation of non-unitary crosstalk in the fiber, MDL can be also present at the level of the optical components in the SDM link.

4.2.3 Case study 2: few mode amplifiers with mode dependent gain (MDG)

4.2.3.1 Channel model

Single fiber span

In this second case study, we intentionally choose to keep the fiber-generated coupling unitary as in [76, 77], neglecting any distributed MDL that could be induced through coupling of the M modes with unguided higher order modes at imperfect slices and fiber micro-bends as considered in [78]. This is done in order to solely focus on the accumulated MDL generated by the inevitable mode dependent gains of FMAs that are hard to cancel due to technological limits: 4-5 dB of MDG in a 3-mode FMA in [179] and 2-3 dB of MDG in a 6-mode FMA in [92].

Our hypothesis of unitary modal coupling is true provided that the fibers are designed to induce very low overall coupling losses [78]. A recent work showed the possibility of realizing low-splice losses of FMF sections [178] and some initial transmission experiments over 30 km [10] to 96 km [83] of FMFs have also shown a negligible fiber-generated MDL. To model the distributed mode coupling, each fiber span consists of K independent fiber sections given by:

$$\mathbf{F}_{span, M \times M} = \prod_{k=1}^K (\mathbf{T}_k \mathbf{R}_k) \quad (4.6)$$

Each section is modeled as a product of a diagonal matrix $\mathbf{T}_k$ with random phase entries uniformly drawn in $[0 : 2\pi]$ to model phase shifts between the propagating modes, and a real orthogonal rotation matrix $\mathbf{R}_k$ representing the distributed modal crosstalk.

In [77], 2×2 rotation matrices are used for two modes LP_{01} and LP_{11} (neglecting its degeneracies a and b) and the rotation angle is drawn from a Gaussian distribution with a standard deviation determining the coupling strength; whereas in [76], the 2×2 unitary crosstalk matrices are generated using the coupled-mode theory (CMT) where the coupling strength depends on the random fluctuation of the core center as well as on the difference between the mode propagation constants. In both cases, the number of modes is limited to 2. However, a general and simple procedure is needed to generate rotation matrices that model different crosstalk levels for any number of modes. To achieve this, the mode coupling angles of the real rotation matrices $\mathbf{R}_k$ will be computed using the crosstalk levels generated at “fictional” displaced cores of two fiber sections. These crosstalk levels are computed by an overlap integral of the electrical field distributions of the propagating modes over the fiber cross section as in [90]:

$$c_{M1-M2} = \iint_A E_{M1}(x, y) E_{M2}^*(x + \Delta x, y + \Delta y) dA \quad (4.7)$$

where $M1$ and $M2 \neq M1$ are two different guided modes. It is important to mention that we are just using overlap integrals at a core misalignment as a mathematical tool to emulate modal coupling.

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

In case a real misalignment was present, the resulting crosstalk matrix would not be unitary as seen in the previous section and in [90]. This emulation technique was motivated by the observation in [180] where authors noted that the crosstalk levels computed by overlap integrals were the same as the ones computed from the CMT that better models the physical origins of coupling. However, the simulation of CMT for a large number of modes is resource-hungry in terms of computing power and less practical than our proposed method.

$M(M-1)/2$ crosstalk levels are needed for each real rotation matrix $\mathbf{R}_k$, corresponding to the number of its rotation angles. For instance, for 3 modes, each crosstalk value $c_{LP_{01}-LP_{11a}}$, $c_{LP_{01}-LP_{11b}}$, $c_{LP_{11a}-LP_{11b}}$ is equivalent to an Euler angle of a rotation matrix given by:

$$\mathbf{R}_{3 \times 3}(\phi, \theta, \psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.8)$$

where $\phi = \arcsin(c_{LP_{11a}-LP_{11b}})$, $\theta = \arcsin(c_{LP_{01}-LP_{11b}})$ and $\psi = \arcsin(c_{LP_{01}-LP_{11a}})$. Computing 6-mode crosstalk matrices is a generalization of the latter procedure where 15 crosstalk values are needed. The misalignments Δx and Δy for each section are drawn from a zero-mean Gaussian distribution and standard deviation σr_c where σ is a percentage of the core radius r_c determining the coupling strength.

Complete multi-span system

The multi-span SDM optical transmission system will be studied with three (respectively six) single-polarization propagating modes: the fundamental mode LP_{01} and the two degeneracies LP_{11a} and LP_{11b} of the LP_{11} mode (respectively the three previously mentioned modes and the LP_{02} , LP_{21a} and LP_{21b} modes). An OFDM signal along with a suitable cyclic prefix is used to modulate each mode that will propagate through a long-haul optical link containing few-mode amplifiers with modal gain offsets. The transmission scheme is shown in Fig. 4.11. The resulting MIMO channel for each subcarrier is given by:

$$\mathbf{Y}_{M \times T} = \mathbf{H}_{M \times M} \mathbf{X}_{M \times T} + \mathbf{N}_{M \times T} = \sqrt{\alpha} \prod_{l=1}^L (\mathbf{P}_l \mathbf{G}_l \mathbf{F}_l) \mathbf{X}_{M \times T} + \mathbf{N}_{M \times T} \quad (4.9)$$

where $\mathbf{X}_{M \times T}$ (resp. $\mathbf{Y}_{M \times T}$) are the emitted (resp. the received) complex symbols on the $M = \{3, 6\}$ modes and during T time slots. $\mathbf{H}_{M \times M}$ is the channel matrix consisting of L independent fiber spans $\mathbf{F}_l$ given by (4.6), followed each by an FMA modeled as a diagonal matrix $\mathbf{G}_l$, as well as a mode scrambler $\mathbf{P}_l$. The gains in $\mathbf{G}_l$ are assigned as follows: the LP_{01} mode has a unit gain and the gain level of the $LP_{\mu\nu}$ mode is given by $\exp(\Delta G_{01-\mu\nu} \ln 10 / 20)$, $\Delta G_{01-\mu\nu}$ being the gain offset in dB. $\mathbf{P}_l$ are random permutation matrices obtained by randomly permuting the rows of an identity matrix $\mathbf{I}_M$, representing perfect mode mixers. α is a normalization factor compensating

the common link loss. Again, modal dispersion is not considered since it does not affect the capacity of the system. Hence, the investigated channel matrix $\mathbf{H}$ is frequency independent or is observed at the level of a single OFDM subcarrier. In reality, $\mathbf{H}$ and in consequence MDL, will be frequency-dependent with a coherence bandwidth that depends on the amount of modal dispersion [74, 91]. As a first approach, we will limit our study to frequency-independent MDL.

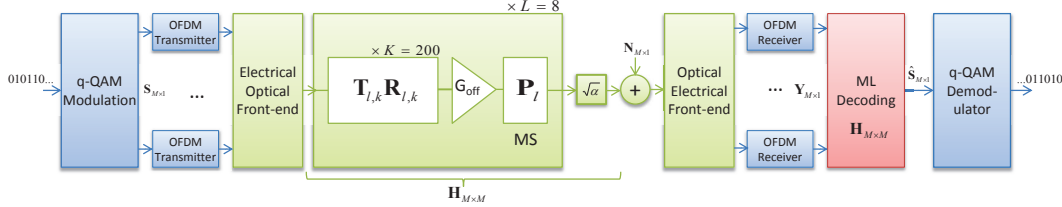


Figure 4.11: Mode-multiplexed OFDM optical transmission system.

Finally, $\mathbf{N}_{M \times T}$ is an additive noise assumed to be zero-mean white Gaussian of variance N_0 per real dimension per mode added at the receiver. In a long-haul optical link, the major noise source is the amplified spontaneous emission (ASE) noise generated at each amplification stage. This noise is distributed and is subject to MDL while it propagates with the signal. Hence, the noise level per mode can vary, resulting in a spatially colored noise. However, it was observed in [74] that the standard variation of the spatial non-whiteness of the noise remains far smaller than the signal power variations induced by MDL as the number of noise sources (amplifiers) in the link increases. Hence, the hypothesis of a white noise is maintained in our model.

Before simulating the channel, we will look into the statistics of MDL in our channel model. To evaluate the MDL under different coupling scenarios, we consider an SDM system with $L = 8$ spans where FMFs have a parabolic index profile with a core radius of $8.7 \mu\text{m}$ supporting 6 modes. Each span consists of $K = 200$ sections. The amplifiers present the following gain offsets $\Delta G_{01-11} = -1.3 \text{ dB}$, $\Delta G_{01-21} = -2 \text{ dB}$ and $\Delta G_{01-02} = -0.2 \text{ dB}$ corresponding to a promising FMA technology presented in [92], and are followed by a mode scrambler. For $K = 200$ sections, three coupling strengths are investigated by drawing misalignments from a Gaussian distribution with σ tuned to 0.6%, 3% and 5% of the core radius r_c to emulate weak, medium and strong mode coupling respectively. 10^6 channel realizations were numerically simulated and the corresponding MDL levels were computed for each scenario, MDL being defined as the ratio in dB between the squares of the highest and the lowest singular values of $\mathbf{H}$. The obtained MDL distributions are shown in Fig. 4.12 for the different coupling levels, with and without the mode scrambling option.

First, we notice that coupling and scrambling significantly reduce the average MDL as well as its variance. With weak-coupling fiber spans, the 2 dB MDL of each amplifier sums up resulting in an accumulated MDL of $8 \times 2 \text{ dB} = 16 \text{ dB}$. At medium coupling, the modes are only

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

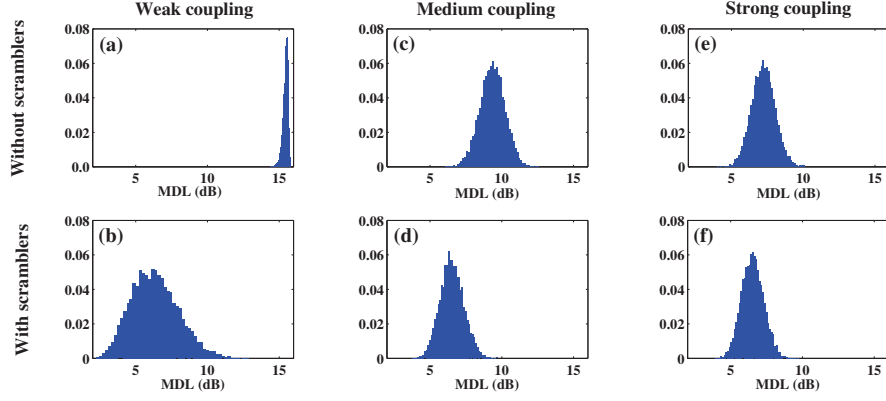


Figure 4.12: Simulated probability distribution functions of MDL in the investigated 6×6 MDM system in different coupling and scrambling scenarios (8 spans, 2 dB of MDL per span).

partially correlated and the average MDL decreases to 10 dB. Strong coupling and mode scrambling reduces the average MDL to 6 dB which is very close to $\sqrt{8} * 2 \text{ dB} = 5.7 \text{ dB}$, the expected accumulated MDL value when full, random coupling occurs between identical MDL sources [74]. Hence, strong coupling in MMF is desirable in order to reduce the accumulated MDL as it does with the modal group delay spread. However, MDL is not completely eliminated. Again, we note that, only in the strong coupling regime, the MDL statistics coincide with the ones obtained from the model suggested by Kahn *et al.*

The impact of MDL on the capacity C of the MDM channel in (4.9) is illustrated in Fig. 4.13 where the cumulative distribution functions (CDF) of the capacity, at $\text{SNR}_{dB} = 10 \log_{10}(E_S/2N_0) = 10 \text{ dB}$ per mode, are given. The CDF of C is computed as explained in section 2.3.3 for a general MIMO channel. The capacity of a 6×6 MDL-free additive white Gaussian noise (AWGN) channel is also shown as a reference.

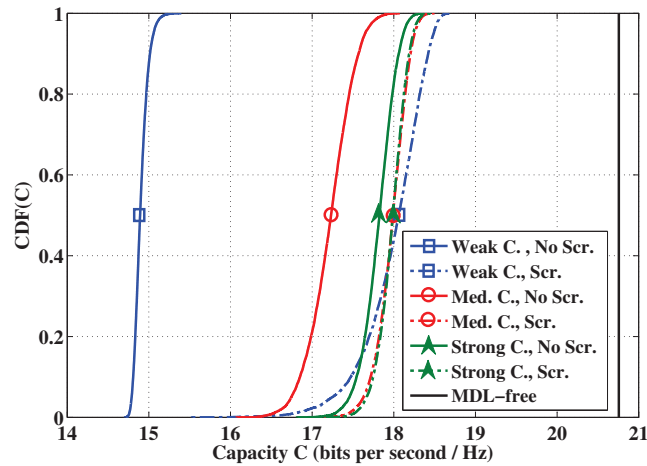


Figure 4.13: CDF of the capacity of the 6×6 SDM system in the different coupling and scrambling scenarios at $\text{SNR} = 10 \text{ dB}$.

4.2.3.2 MDL-induced penalties

After presenting the second SDM channel model in Fig. 4.11, we measure the penalties induced by MDG in a 3-mode (resp. 6-mode) system where graded-index fibers with a parabolic index profile are installed, with a core radius $r_c = 6 \mu\text{m}$ (resp. $r_c = 8.7 \mu\text{m}$). For the 6-mode system, the link parameters are the same as the one used in the previous section ($L = 8$ spans, each made of $K = 200$ sections with three modal coupling levels in the fibers, and the given gain offsets at FMAs). For the 3-mode system, a gain offset of $\Delta G_{01-11} = -1.3$ dB is considered at each FMA.

At the transmitter, a vector of 4-QAM symbols $\mathbf{s}_{m=1:M}$ of unit energy $E_S = 1$ is sent over each mode, providing a spectral efficiency of six bits (resp. 12 bits) per time slot. The performance in terms of average BER curves versus the symbol signal-to-noise ratio $E_S/2N_0$ per mode is measured through Monte-Carlo simulations. A minimum of 100 bit errors and a minimum of 1000 channel realizations are registered per simulation point. Three coupling scenarios with and without the scrambling option at the FMAs are presented in Fig. 4.14. At the receiver, in all scenarios, the data symbols are retrieved using an ML sphere decoder.

The first column in Fig. 4.14 shows the results for a 3-mode MDM system. From the square-marked curves corresponding to NC without mode scramblers, we notice that the SNR penalty at $\text{BER} = 10^{-3}$ induced by MDL decreases from 4.2 dB for weak coupling, to 1.2 dB for medium coupling and to 0.5 dB for strong coupling. Adding mode scramblers at FMAs (triangle marked curves), reduces these penalties to 1.5 dB for weak coupling and 0.4 dB for medium coupling while it has no effect in strong coupling regime because the modes are already fully coupled in the fiber.

The same observations can be made for the 6-mode MDM system (second column in Fig. 4.14). Without scramblers, the SNR penalty at $\text{BER} = 10^{-3}$ decreases from more than 6 dB for weak coupling, to 1.2 dB for medium coupling and to 0.4 dB for strong coupling. The insertion of mode scramblers reduces these penalties to 1.7 dB for weak coupling and 0.4 dB for medium coupling. Again, scrambling has no effect in the strong coupling regime because the modes are fully coupled in the fiber and the MDL cannot be further reduced by the scramblers as can be seen from the MDL distributions in Fig. 4.12. Modal coupling and scrambling reduce the overall MDL and thus, enhance the performance of MDL-impaired MDM schemes, which was already observed in previous works [2, 94].

4.3 ST coding for various MDL-impaired MDM channel models

In order to reduce MDL, we have seen that stringent requirements must be imposed on splice losses and modal dependencies of the optical components. Despite the use of optical solutions to

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

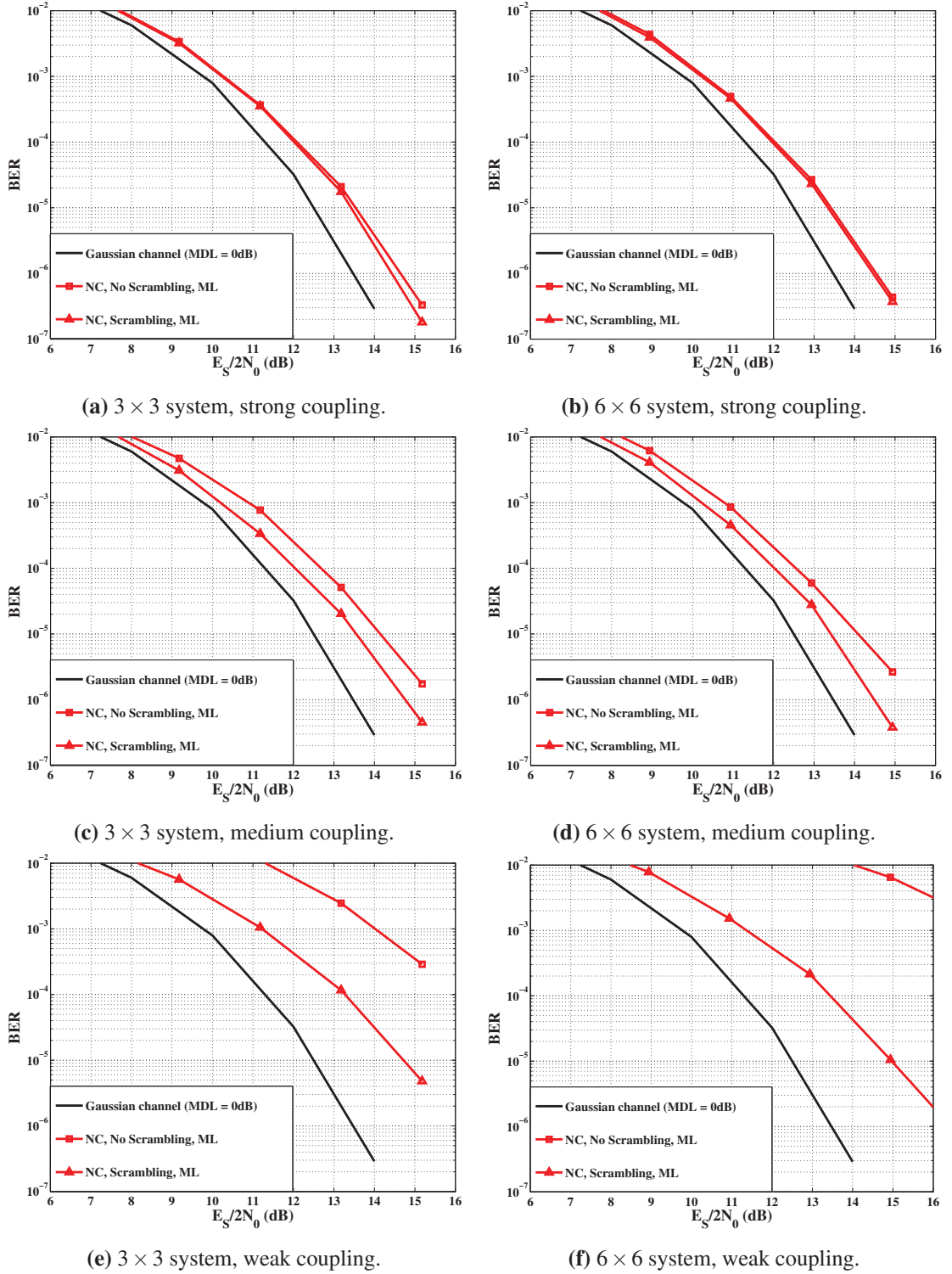


Figure 4.14: Performance in terms of average BER versus SNR of 3×3 (a,c,e) and 6×6 (b,d,f) MDM systems obtained through Monte Carlo simulations (8 spans of graded-index parabolic profile fibers with a core radius $r_c = 6 \mu\text{m}$ (resp. $r_c = 8.7 \mu\text{m}$) and a numerical aperture $NA = 0.205$ at a wavelength $\lambda = 1550 \text{ nm}$, FMAs with a maximum modal gain offset of 2 dB).

reduce MDL (mode scrambling and strong coupling), this effect is not suppressed from the optical link. These observations urge the search for more sophisticated MIMO tools to mitigate MDL.

Conventional optical MIMO schemes are used for simple multiplexing which consists, in the SDM case, in sending, at the transmitter, a vector $\mathbf{s}_{M \times 1}$ of independent q-QAM symbols on M modes at a single time slot. However, we can benefit from the two degrees of freedom “space” and “time” of the SDM MIMO scheme to insert multiple copies of a data symbol over different modes at different time slots. Then, at the receiver, the copies are exploited to enhance the performance since the data symbols would have experienced various channel states and the modal loss disparities would be further reduced. In the following, STBCs are tested in order to mitigate MDL. We investigate the performance of ST-coded 3-mode and 6-mode SDM schemes where MDL is introduced first by imperfections in the waveguide, and second by gain offsets at FMAs. The general transmission scheme is presented in Fig. 4.15.

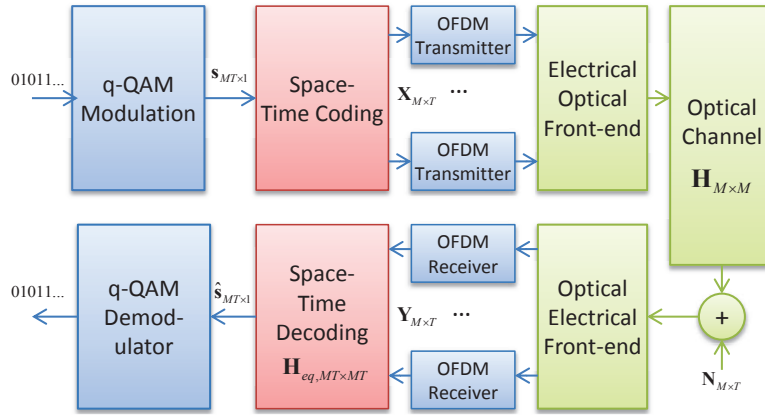


Figure 4.15: Mode-multiplexed Space-Time coded OFDM optical transmission system.

4.3.1 Case study 1: fiber misalignments and micro-bends

For our first case study on fiber misalignments and micro-bends, we consider the 3-mode system simulated in section 4.2.2. Among STBCs, we will choose 3×3 codes that fulfill the following requirements: place each data symbol on a different mode at each time slot in order to average the losses, avoid any redundancy by choosing a full-rate code meaning that 9 q-QAM information symbols are sent in each codeword, and finally transmit a uniform average energy per mode. TAST codes answer these requirements and are known for their generality and performance [148]. The TAST code family is presented in chapter 2 and the 3×3 TAST codeword is given by:

$$\mathbf{X}_T = \frac{1}{\sqrt{3}} \begin{bmatrix} s_1 + \theta s_2 + \theta^2 s_3 & \phi^2(s_4 + j\theta s_5 + j^2\theta^2 s_6) & \phi(s_7 + j^2\theta s_8 + j\theta^2 s_9) \\ \phi(s_7 + \theta s_8 + \theta^2 s_9) & s_1 + j\theta s_2 + j^2\theta^2 s_3 & \phi^2(s_4 + j^2\theta s_5 + j\theta^2 s_6) \\ \phi^2(s_4 + \theta s_5 + \theta^2 s_6) & \phi(s_7 + j\theta s_8 + j^2\theta^2 s_9) & s_1 + j^2\theta s_2 + j\theta^2 s_3 \end{bmatrix} \quad (4.10)$$

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

where $\phi = \exp(i\pi/36)$, $j = \exp(i2\pi/3)$, $\theta = \exp(i\pi/9)$ and $s_{k=1:9}$ are the modulated symbols of unit energy. ϕ and θ are chosen to maximize the coding and diversity gains over a wireless Rayleigh fading 3×3 multi-antenna channel [148]. 9 different symbols are sent on 3 time slots in each codeword, achieving a rate of 3 symbols/cu. Another way of representing the code consists in vectorizing, column-wise, (4.10) into the following form:

$$\begin{aligned} \text{vec}_{\mathbb{C}}(\mathbf{X}_{\mathcal{T}}) &= \mathbf{M}_{G,MT \times MT} \mathbf{s}_{MT \times 1} \\ &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & \theta & \theta^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \phi\theta & \phi\theta^2 \\ 0 & 0 & 0 & \phi^2 & \phi^2\theta & \phi^2\theta^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi^2 & \phi^2j\theta & \phi^2j^2\theta^2 & 0 & 0 & 0 \\ 1 & j\theta & j^2\theta^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \phi j\theta & \phi j^2\theta^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \phi j^2\theta & \phi j\theta^2 \\ 0 & 0 & 0 & \phi^2 & \phi^2j^2\theta & \phi^2j\theta^2 & 0 & 0 & 0 \\ 1 & j^2\theta & j\theta^2 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \end{bmatrix} \quad (4.11) \end{aligned}$$

where $\mathbf{M}_G$ is the generator matrix of the code and is unitary ($\mathbf{M}_G \mathbf{M}_G^\dagger = \mathbf{I}_{9 \times 9}$). Hence, we do not increase the energy of the new transmitted symbols after encoding the information symbols.

Assuming that the channel matrix $\mathbf{H}$ is known (or perfectly estimated at the receiver) and constant during T time slots, and that all emitted codewords $\mathbf{X}$ are equiprobable, the optimal detection scheme is the ML criterion that will be implemented at the receiver side using a sphere decoder. An equivalent channel is defined to decode the ST-coded symbols by using a vectorized form explicitly showing the generator matrix of the ST code as explained in section 2.4.4.

Similar Monte-Carlo simulations as the one in the previous sections are run. When the TAST code is used alone (circle-marked red curves in Fig. 4.16a, 4.16b and 4.16c), the achieved performance is close to the one obtained with a scrambling period $K_{scr} = 16$. We absorb all the penalty at $2\%r_c$, and are left with a penalty of 0.5 dB at $3\%r_c$ and 1 dB at $4\%r_c$ which is equivalent to an important coding gain of 3 dB. Combining the code with the scramblers produces no penalty in all cases and not only at a BER of 10^{-3} but also for any achieved BER. This observation is valid for $K_{scr} = 1$ and $K_{scr} = 16$, hence a longer scrambling period K_{scr} would be needed if the TAST code is used.

The choice of the scrambling periods K_{scr} is purely conceptual and optimistic comparing to the number of mode scramblers that may be installed in a real mode-multiplexed optical link. Besides, the techniques proposed in [94] to achieve mode scrambling have imperfections such as important insertion losses and are far from being equivalent to perfect permutations as modeled in our simulation. In a real optical link, mode scramblers would be installed mainly at the optical amplification stage which limits the number of scramblers in the link. On the other hand, the above results show that ST coding is an interesting candidate for MDL mitigation and has the advantage

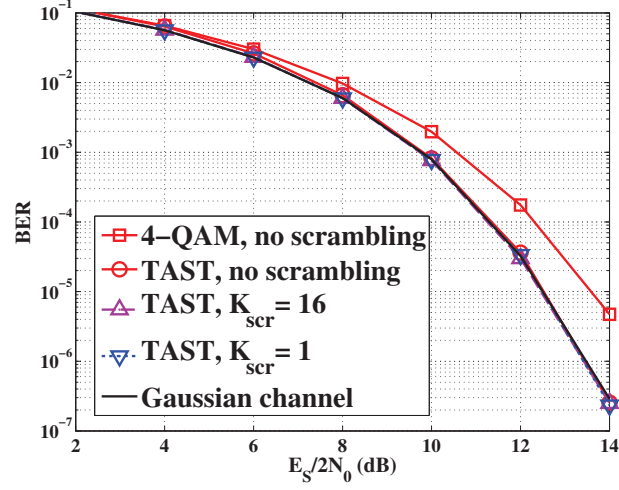
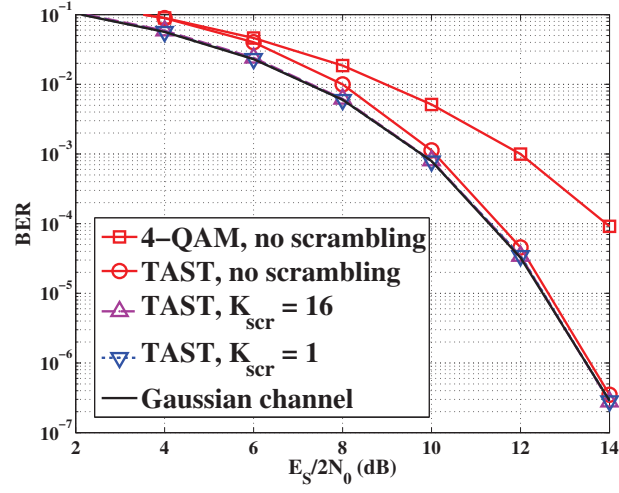
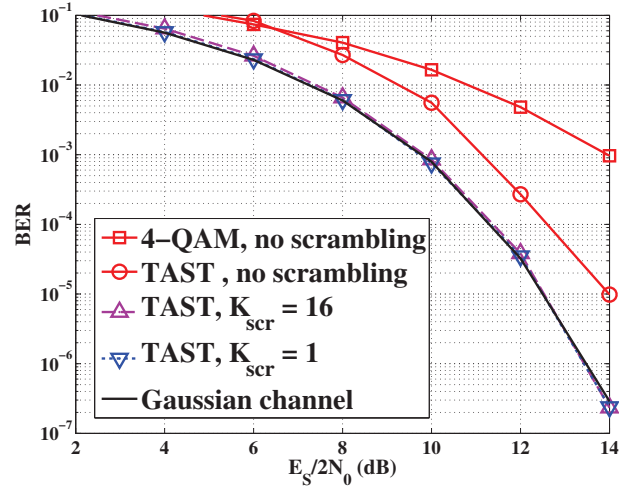

 (a) $\sigma_{x,y} = 2\%r_c, ST$ coding

 (b) $\sigma_{x,y} = 3\%r_c, ST$ coding

 (c) $\sigma_{x,y} = 4\%r_c, ST$ coding

Figure 4.16: BER curves of the 3×3 TAST-coded SDM system ($K = 300$) with a spectral efficiency of 6 bits/c.u. and using ML decoding.

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

of being a software based solution implemented at the transmitter. Hence, it can be adopted as an alternative to mode scrambling or as a complementary solution to limit the number of installed scramblers in the optical link. In addition, ST codes exist for $M > 3$ and thus, can be extended to SDM systems with more than 3 modes.

In order to test the potential of the TAST code in mitigating higher MDL levels, we simulate the SDM channel with the same parameters as before and an increased misalignment std $\sigma_{x,y} = 5\%$. The performance in terms of BER is given in Fig. 4.17a along with the corresponding MDL statistics for each scrambling scenario in Fig. 4.17b.

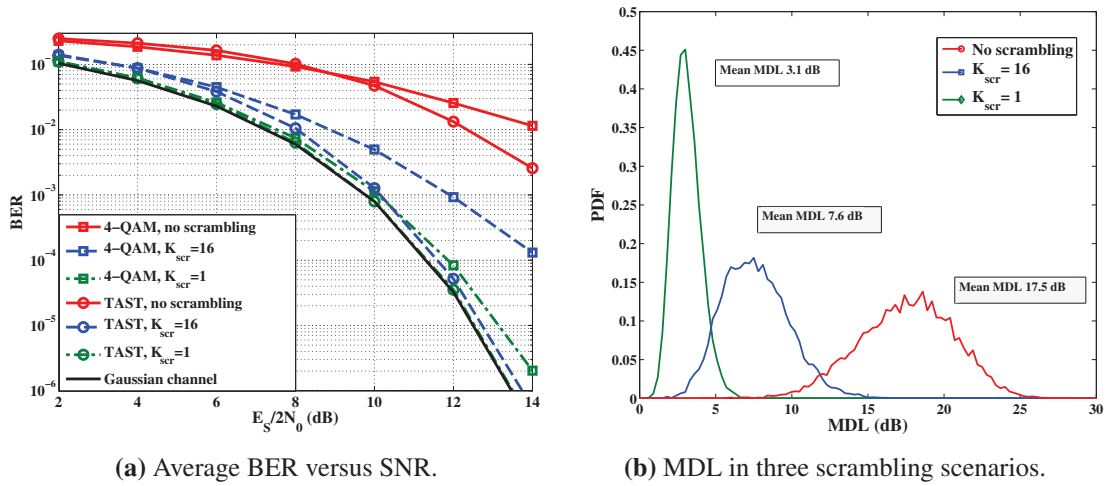


Figure 4.17: BER curves and MDL statistics of the 3×3 TAST-coded SDM system ($K = 300$) with $\sigma_{x,y} = 5\%r_c$ and using ML decoding.

Without scrambling, MDL levels are very high (up to 25 dB) and the coding gain is limited. We notice that the TAST code needs to be complemented with more scramblers in order to fully absorb the MDL-induced penalties. With a scrambling period of $K_{scr} = 16$, a residual SNR penalty of 0.3 dB is observed at $\text{BER} = 10^{-3}$. In this scenario, the average MDL is equal to 7.6 dB.

4.3.2 Case study 2: few mode amplifiers with mode dependent gain (MDG)

4.3.2.1 Performance analysis of ST codes

For the second case study, we consider 3-mode and 6-mode SDM systems impaired by MDG arising at few mode amplifiers using the same parameters as before, and test the ability of 3×3 and 6×6 TAST codes to eliminate the accumulated MDL. The 6×6 TAST codeword is given by:

$$\mathbf{X}_{\mathcal{T}} = \frac{1}{\sqrt{6}} \begin{bmatrix} f_1(\mathbf{s}_1) & \phi^{\frac{5}{6}} f_2(\mathbf{s}_6) & \phi^{\frac{4}{6}} f_3(\mathbf{s}_5) & \phi^{\frac{3}{6}} f_4(\mathbf{s}_4) & \phi^{\frac{2}{6}} f_5(\mathbf{s}_3) & \phi^{\frac{1}{6}} f_6(\mathbf{s}_2) \\ \phi^{\frac{1}{6}} f_1(\mathbf{s}_2) & f_2(\mathbf{s}_1) & \phi^{\frac{5}{6}} f_3(\mathbf{s}_6) & \phi^{\frac{4}{6}} f_4(\mathbf{s}_5) & \phi^{\frac{3}{6}} f_5(\mathbf{s}_4) & \phi^{\frac{2}{6}} f_6(\mathbf{s}_3) \\ \phi^{\frac{2}{6}} f_1(\mathbf{s}_3) & \phi^{\frac{1}{6}} f_2(\mathbf{s}_2) & f_3(\mathbf{s}_1) & \phi^{\frac{5}{6}} f_4(\mathbf{s}_6) & \phi^{\frac{4}{6}} f_5(\mathbf{s}_5) & \phi^{\frac{3}{6}} f_6(\mathbf{s}_4) \\ \phi^{\frac{3}{6}} f_1(\mathbf{s}_4) & \phi^{\frac{2}{6}} f_2(\mathbf{s}_3) & \phi^{\frac{1}{6}} f_3(\mathbf{s}_2) & f_4(\mathbf{s}_1) & \phi^{\frac{5}{6}} f_5(\mathbf{s}_6) & \phi^{\frac{4}{6}} f_6(\mathbf{s}_5) \\ \phi^{\frac{4}{6}} f_1(\mathbf{s}_5) & \phi^{\frac{3}{6}} f_2(\mathbf{s}_4) & \phi^{\frac{2}{6}} f_3(\mathbf{s}_3) & \phi^{\frac{1}{6}} f_4(\mathbf{s}_2) & f_5(\mathbf{s}_1) & \phi^{\frac{5}{6}} f_6(\mathbf{s}_6) \\ \phi^{\frac{5}{6}} f_1(\mathbf{s}_6) & \phi^{\frac{4}{6}} f_2(\mathbf{s}_5) & \phi^{\frac{3}{6}} f_3(\mathbf{s}_4) & \phi^{\frac{2}{6}} f_4(\mathbf{s}_3) & \phi^{\frac{1}{6}} f_5(\mathbf{s}_2) & f_6(\mathbf{s}_1) \end{bmatrix} \quad (4.12)$$

where $\phi = \exp(i\pi/12)$, $\mathbf{s}_{1:6}$ a vector of 6 data symbols, $f_n(\mathbf{x}) = \sum_{k=1:6} x_k (j^{n-1} \theta)^{k-1}$ with $j = \exp(i2\pi/6)$, $\theta = \exp(i\pi/18)$. ϕ and θ are chosen to maximize the coding and diversity gains over a wireless Rayleigh fading 6×6 multi-antenna channel. 36 symbols are sent on 6 time slots in each codeword, achieving a full rate of 6 symbols/cu. Its generator matrix, which is also unitary, can be found in [148].

The performance in terms of average BER curves versus $E_S/2N_0$ per mode of the ST-coded schemes are measured through Monte-Carlo simulations. Three coupling scenarios with and without the scrambling option at the FMAs are presented in Fig. 4.18. Assuming that the channel matrix $\mathbf{H}$ is known (or perfectly estimated at the receiver) and constant during T time slots, and that all emitted codewords $\mathbf{X}$ are equiprobable, the optimal detection scheme is the ML criterion that will be implemented at the receiver side using a sphere decoder. An equivalent channel is defined to decode the ST-coded symbols by using a vectorized form explicitly showing the generator matrix of the ST code as explained in section 2.4.4.

When the 3×3 TAST code is used alone (square-marked dashed curves in Fig. 4.18), it outperforms the scrambling solution in the scenario of weak coupling, reducing the SNR penalty at $\text{BER} = 10^{-3}$ to 1 dB for a 3-mode system, which is equivalent to a coding gain of 3 dB. Furthermore, the TAST code absorbs all the MDL-induced penalty in the medium and strong coupling scenarios. Similar reductions of penalties are observed with the 6×6 TAST code for the different coupling levels. Finally, combining the ST code with mode scrambling results in no penalty in all 3-mode and 6-mode schemes, and not only at a BER of 10^{-3} but for any given BER. These results prove the efficiency of ST coding solutions in mitigating MDL by averaging the losses experienced by the mode multiplexed data symbols, making it an interesting DSP solution for SDM systems.

From the obtained results in the two case studies, we conclude that ST codes can be adopted as an alternative to mode scrambling or as a complementary solution to limit the number of mode scramblers in the optical link depending on the coupling strength in the installed FMFs. ST coding can relax modal gain offset requirements in optical components or equivalently improve the reach of the SDM system through the offered coding gains.

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

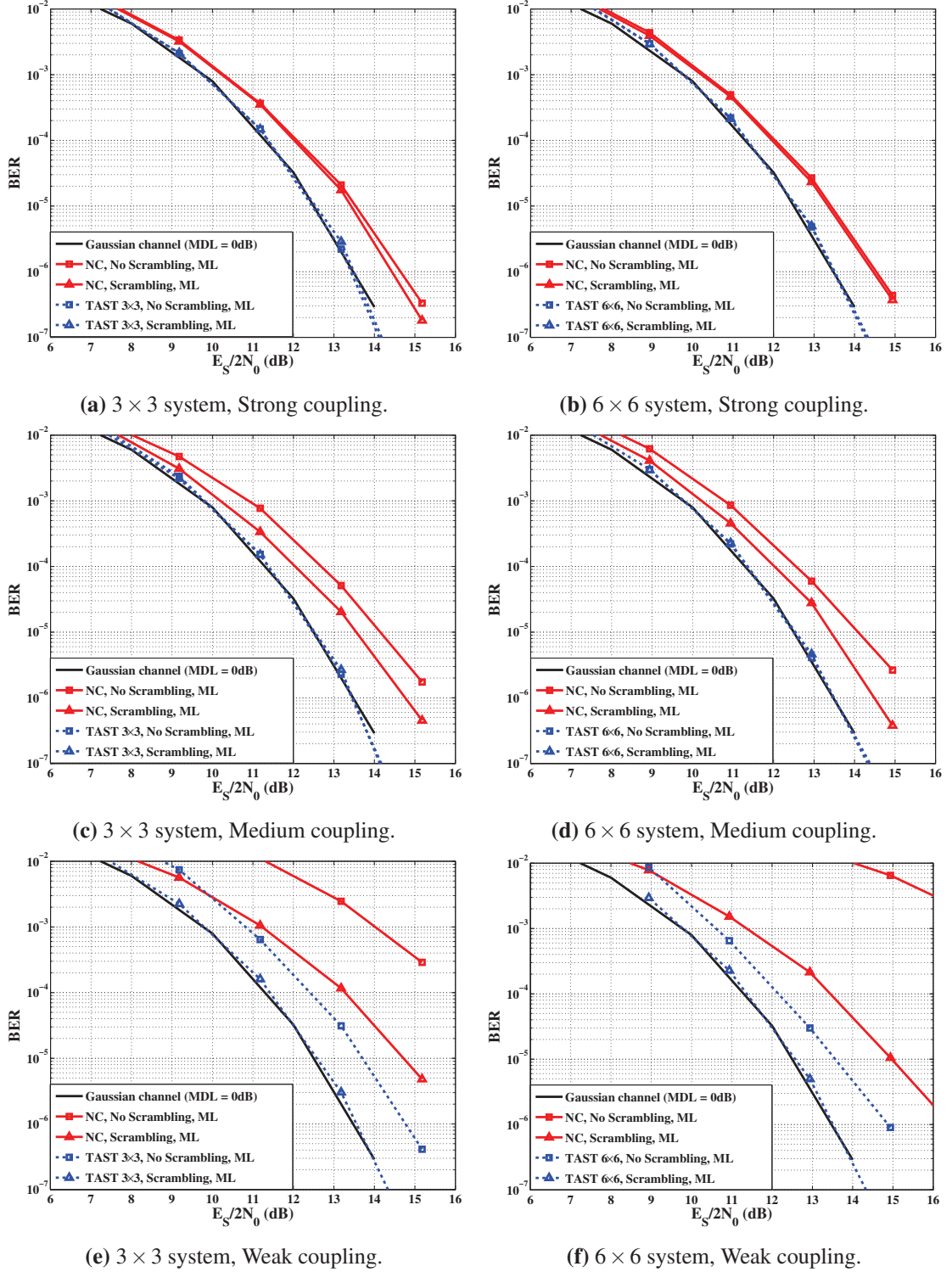


Figure 4.18: Performance in terms of average BER versus SNR of 3×3 (a,c,e) and 6×6 (b,d,f) ST-coded MDM systems obtained through Monte Carlo simulations (8 spans of graded-index parabolic profile fibers with a core radius $r_c = 6 \mu\text{m}$ (resp. $r_c = 8.7 \mu\text{m}$) and a numerical aperture $NA = 0.205$ at a wavelength $\lambda = 1550 \text{ nm}$, FMAs with a maximum modal gain offset of 2 dB).

We have seen that full absorption of high MDL levels was possible using ST codes along with mode scramblers at each amplification stage. On the other hand, we also observed that modal coupling in installed fibers is able to reduce the accumulated MDL. However, this reduction is lower bounded by an average MDL that grows as the square root of the number of MDL sources (FMAs in our case) in the link as explained in [64, 74]. This minimum average value can be reached, in a system of a given length, using solely fibers with inherently strong coupling as seen from the MDL statistics in Fig. 4.12 or with regular mode scrambling when there is not enough coupling in the fiber. In weak coupling, mode scramblers are needed at each amplification stage. However, for intermediate coupling strengths, a lower number of scramblers might be installed.

4.3.2.2 Optimal scrambling for MDM systems

The placement and number of scramblers required for an optimal reduction of MDL are important parameters for the design of long-haul SDM systems because real mode scramblers are not as perfect as modeled in the numerical studies (perfect permutation matrices). Originally, mode scramblers were used to create a uniform modal distribution for purposes of laboratory, industrial or field measurements of the bandwidth of multi-mode fibers. Practical implementations of mode scramblers include a concatenation of fibers of different index profiles scramblers [181] or passing a fiber through a series of small bends [182] or gratings [80]. These versions of mode scramblers are simple but not easily reproducible and have non-negligible insertion losses. With the rise of interest in SDM, mode scramblers are being considered for long-haul telecommunications. We have seen their benefits in the previous sections. Apart from the mentioned implementations, mode scramblers can also be realized using mode converters with free-space optical components that also have non-negligible crosstalk and insertion losses. Hence, the numbers of mode scramblers in the optical link must be limited.

In order to find the optimal scrambling map for the studied SDM system, we define the ratio $r = N_{scr}/L$ where N_{scr} denotes the number of scramblers per L fiber spans, with $0 \leq r \leq 1$ allowing at most for a single mode scrambler per span. Considering the parameters of the 6-mode system in section 4.2.3, we draw the evolution of the average accumulated MDL with the number of spans L for different coupling strengths (obtained by varying the std of the fictional core displacements σ) and in the absence of mode scramblers. The resulting curves are drawn in Fig. 4.19 along with the lower bound $2\sqrt{L}$ where 2 dB stands for the MDL per FMA.

Then, for each value of σ and number of fiber spans L , we vary the ratio of mode scramblers r in the link and determine the minimum ratio r_{opt} that leads to an average MDL within 1 dB of its lower bound $2\sqrt{L}$ dB. For a small number of spans of weak-coupling fibers, the optimal MDL reduction cannot be reached because the modes are not yet fully coupled as we can see from Fig. 4.12 in the weak coupling scenario with mode scramblers at each FMA ($r = 1$). In this case,

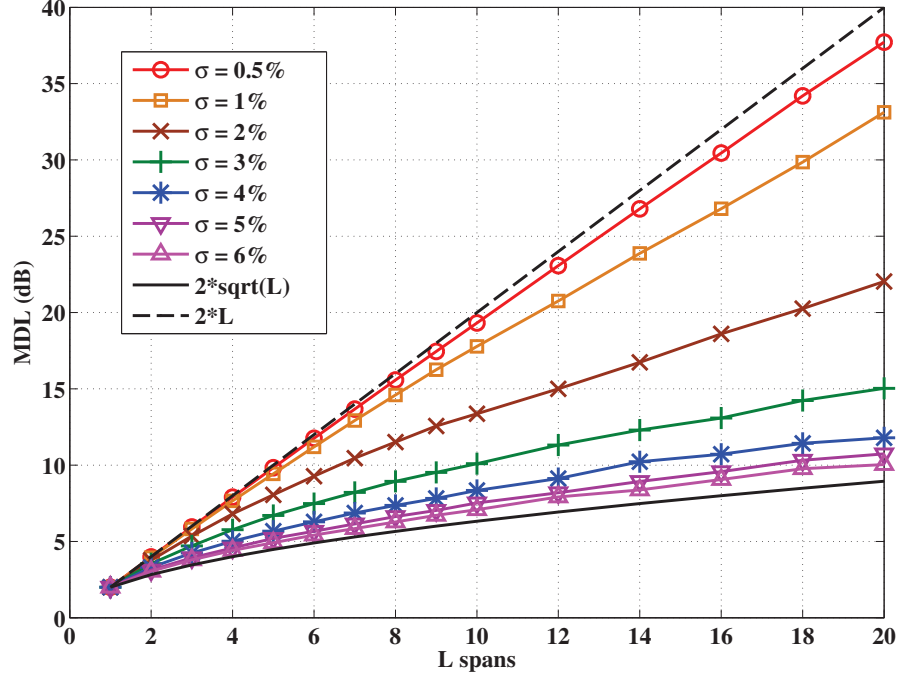


Figure 4.19: Evolution of the average accumulated MDL with the number of spans L for various coupling strengths in absence of mode scramblers (2 dB MDL per span).

a longer system should be considered to reach the optimal MDL reduction that we will call the asymptotic regime in which the modes are fully coupled. The obtained values of r_{opt} are reported in Fig. 4.20 for different σ and L . The following discrete values of r were tested for each coupling strength: 0, $1/10$, $1/5$, $1/4$, $1/3$, $1/2$, $2/3$, $3/4$, $4/5$, 1.

We can observe that, for $1\% \leq \sigma \leq 6\%$, the number of required scramblers decreases with the length of the transmission system until it reaches an asymptotic value when the fully-coupled regime is reached. Fig. 4.21 shows the asymptotic values of r_{opt} obtained for different coupling strengths as well as a linear interpolation for intermediate values. We see that for $\sigma \leq 1\%$, a mode scrambler is needed at each span and beyond 5%, mode scramblers are no longer needed.

The efficiency of ST coding in mitigating MDL paves the way for further investigation of ST coding schemes in SDM systems with a higher number of modes M or with polarization multiplexing. In chapter 3, we showed that the performance enhancement offered by the codes in PDL-impaired systems obeyed different criteria than those in wireless systems. An analytic expression of the minimum distance between the possible emitted symbols, after propagating in MDL-impaired systems, was given in chapter 3. ST-coded schemes exhibited a larger minimum distance compared to uncoded-schemes which explained the enhanced performance. We conjecture that the same behavior is observed in the MDL-impaired systems and leave the theoretical investigation of the observed coding gains for future studies while we focus, in the next section, on the complexity and scale-up possibilities of ST coding solutions.

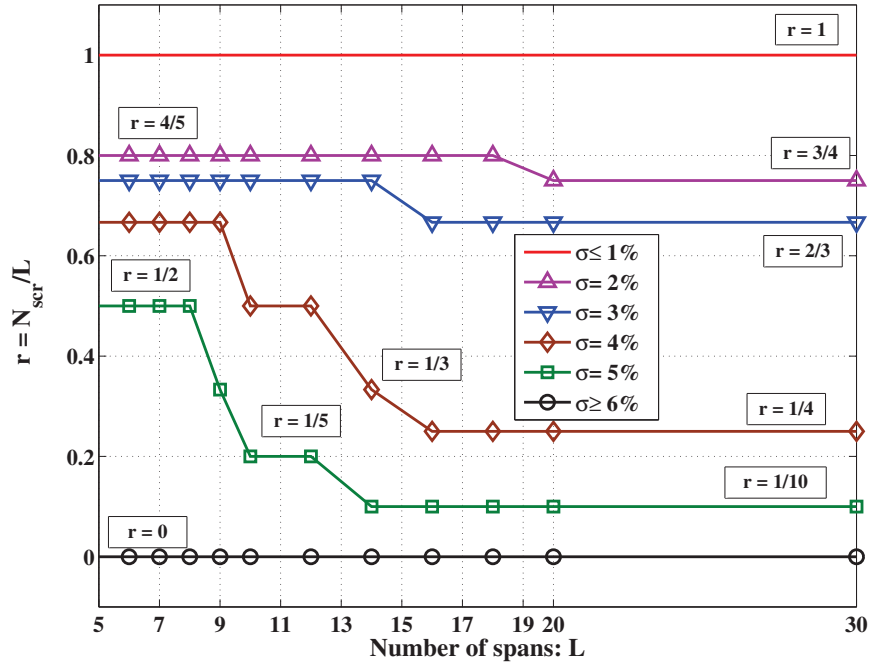


Figure 4.20: Ratio of mode scramblers required to optimally reduce MDL for various coupling strengths determined by σ .

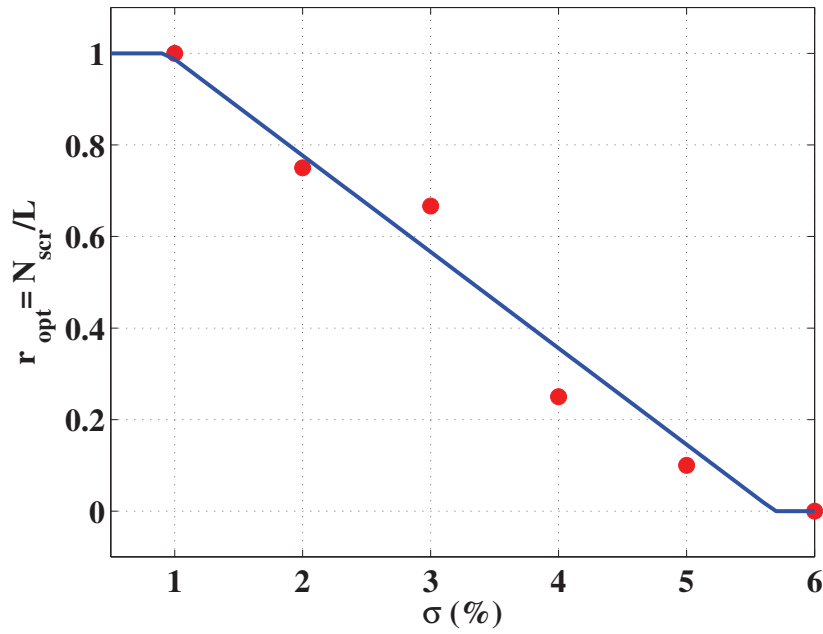


Figure 4.21: Ratio of mode scramblers required to optimally reduce MDL for various coupling strengths determined by σ in the asymptotic regime.

4.4 Scalability and complexity issues of ST coding

Full-rate ST coding solutions exist for MIMO systems of different dimensions and thus can be generalized to larger $M \times M$ SDM systems. However, a down side of ST coding is their decoding complexity that increases fast with the size of the MIMO system. The complexity of an exhaustive search ML decoder increases exponentially with M . The norm in (2.49) must be computed for all possible combinations of the emitted symbols. In the case of a full-rate square ST code with q -QAM symbols, this leads to q^{M^2} norm computations. The complexity of the exhaustive search increases with the constellation size and exponentially with the number of modes.

Lower-complexity decoders implementing the ML criterion such as the sphere decoder used in this chapter exhibit a complexity that is independent of the constellation size q . An exact number of operations cannot be obtained for a sphere decoder but if the initial radius of the sphere decoder is judiciously chosen in function of the SNR and the singular values of the channel, the complexity of the search algorithm is approximated by $\mathcal{O}((MT)^6)$ operations [156]. This gives $\mathcal{O}(M^6)$ operations for uncoded spatial multiplexing and $\mathcal{O}(M^{12})$ for a full-rate square ST code.

In table 4.2, we show the complexity of an ML-decoded $N \times N$ MIMO system where $N = MT$ as well as numerical applications for spatially multiplexed and ST-coded 3-mode and 6-mode schemes using 4-QAM symbols. We computed a rough estimation of this complexity in flops that we define as the number of required complex scalar multiplications in order to decode a transmitted symbol vector $\mathbf{s}_{N \times 1}$. The $N^2 + N$ term in the number of flops of the exhaustive search is the number of scalar complex multiplications needed for one norm computation.

Table 4.2: Comparison of the complexity of different ML decoders of $N \times N$ MIMO systems.

	Exhaustive search	Sphere decoder
$N \times N$ MIMO	$q^N(N^2 + N)$	$\approx N^6$
3-mode (No coding, $N = 3$)	768	729
TAST 3×3 ($N = 9$)	2.36×10^7	5.31×10^5
6-mode (No coding, $N = 6$)	1.72×10^5	4.67×10^4
TAST 6×6 ($N = 36$)	6.3×10^{24}	2.18×10^9

While the complexities of the exhaustive search and the sphere decoder are comparable for the 3×3 uncoded system, the complexity advantage of the sphere decoder becomes more perceptible with an increasing N . For the TAST-coded 6×6 system, the complexity of the sphere decoder is 15 order of magnitudes lower than the exhaustive search. Despite this important complexity reduction, ML decoding of the considered ST schemes using a sphere decoder will still require a large number of operations for larger SDM systems, turning ST codes into a prohibitive MIMO solution. Therefore, in the next chapter, we suggest and analyze two possible variants of ST coding

solutions that trade a portion of the optimal coding gains observed in this chapter for a reduction in complexity and a better scalability.

Summary

In this chapter, we have shown through numerical simulations that ST coding is a promising solution for MDL mitigation in SDM systems. It can be used as a standalone solution or to complement other optical components-based solutions such as mode scrambling depending on the coupling strength in the installed fibers. The ML performance of the applied codes was investigated, showing a complete mitigation of MDL levels up to 10 dB in a 6-mode system, relaxing thus the gain offset requirements in optical components such as FMAs and helping to improve the reach of the transmission system. The requirements in terms of coupling strengths and mode scrambling in the optical link were also investigated in order to find the configurations that minimize MDL to its lowest value, letting the ST code manage the residual MDL. These observations pave the way for further investigation of the capabilities of ST coding such as the maximum loss disparities that can be mitigated, allowing to further reduce the number of required scramblers (r_{opt}), as well as the study of SDM optical links having various MDL sources (fiber imperfections + MDG), a higher number of modes M or polarization-multiplexed SDM channels where PDL can be added as another performance-limiting effect.

4. ST CODING FOR MDL MITIGATION IN MODE DIVISION MULTIPLEXED (MDM) SYSTEMS

Chapter 5

Low-Complexity Schemes for Large Optical MIMO Systems

We have demonstrated in the previous chapters the ability of MIMO digital signal processing using ST coding at the transmitter and ML decoding at the receiver to mitigate non-unitary channel impairments in PDM and MDM optical transmission systems. Implementations of ST coding solutions in PDM systems can be realized today, preferably over the OFDM format that guarantees a low-complexity decoding of ST schemes with its dispersion management insured through the low-rate orthogonal subcarriers and the cyclic prefix. The situation is different for MDM systems especially when we consider a large number of modes $M \geq 6$ in order to significantly increase the capacity. Indeed, the capacity of any suggested MDM transmission system has to exceed by a factor of at least 10 the one of a PDM single-mode transmission to be able to envisage its deployment. As mentioned in chapter 2, large differential delays (DMGD) between modes would require $4M^2$ adaptive time-domain filters larger than the maximal delay (also called memory of the channel) in order to fully remove the induced temporal and spatial inter-symbol-interference.

Among the dispersion management techniques in MDM systems, few-mode fibers (FMFs) supporting the propagation of 3 and 6 spatial modes were designed with low DMGD values and the OFDM modulation format was shown to be among the lowest-complexity schemes with regard to the suppression of inter-symbol interference and equalization at the receiver [86]. However, the reach of MDM-OFDM systems will be limited by the length of the used cyclic prefix (CP) that is required to be small in order to keep a low overhead, and hence avoid an important reduction of the spectral efficiency of the transmission system. Increasing the number of subcarriers to limit the overhead would not be a viable solution because the time-domain OFDM symbols will become longer and hence more sensitive to phase noise and frequency offsets. Another way to manage large dispersion delays consists in setting the size of the CP to a certain maximum allowable overhead, and managing the residual ISI using a front-end time-domain equalizer to shorten

the channel impulse response experienced by the OFDM receiver, a technique known as channel shortening [136].

Moreover, MCF-based systems may be envisaged with closely packed cores where each core can be multi-mode in order to provide high density SDM solutions [1]. The highest number of SDM channels reported so far is 36 where fibers with 12 cores were used with 3 modes per core [176]. Crosstalk will also be inevitable in these SDM systems and should be addressed. However, the complexity of the suggested ML-decoded ST schemes can easily become a hurdle that needs to be overcome in order to scale up these solutions to larger SDM systems.

Indeed, the realization of coherent SDM receivers with advanced DSP is a real challenge. While we used simulation tools and offline processing on experimental frames, sampled by an oscilloscope, that might take few seconds to decode a coded PDM-OFDM frame and even more than one minute for coded SDM systems on a personal computer, a real receiver is required to perform the same processing in real time. The received signals need to be digitized by analog-to-digital converters (ADCs) after optical de-multiplexing and down-conversion to the electrical domain. The latest high-speed ADC prototypes presented in recent conferences on electronics show 90 GS/s speeds and an effective number of bits (ENoB) of 8 bits per sample [102]. Besides, the DSP algorithms have to be implemented on an electronic chip (ASIC, FPGA) that operates at a limited rate and would most probably require many chips processing the incoming data in parallel. The more complex the processing is, the larger the number of required logic gates in the integrated circuit and the more power will be dissipated by the transceivers. Hence, complexity and energy consumption of Terabit-achieving schemes are required to be low to set us on the path towards a commercially available transmission system. We investigate in this chapter the performance of three low-complexity SDM-MIMO schemes that can be easily scaled to large optical MIMO systems. We assume that low DMGD values can be maintained and managed through cyclic-prefix insertion and channel shortening techniques and address the non-unitary MDL effect.

The first two strategies consist in trading a part of the optimal coding gain of the ST-coded SDM schemes with a reduced complexity and a potential for scalability. Both approaches address the decoding given that the encoding block of an STBC scheme reduces to a simple product of a generator matrix multiplied by a symbol vector. The third strategy is based on mode selection. Accordingly, we suggest in this chapter:

- replacing the optimal ML decoder with sub-optimal low-complexity threshold-decision decoders and testing the performance of a Zero-Forcing (ZF) decoder, a ZF with decision feedback equalizer (ZF-DFE) and an MMSE decoder. This strategy is presented in section 5.1.
- replacing the square $M \times M$ code with a multi-block coding over less time slots $T < M$ in order to reduce the complexity of the ML sphere decoder by shortening the dimension of

the codeword dictionary. Hence, less symbols need to be decoded at once. This will be investigated in section 5.2.

- looking at the particularities of the SDM systems impaired by MDL and determining an appropriate set of modes over which the information can be multiplexed. This technique is referred to as mode selection and will be explained in section 5.3.

5.1 Low-complexity sub-optimal decoding of ST codes

The first solution consists in using sub-optimal decoding of ST-coded schemes. The investigated decoders are the three presented in chapter 2: ZF, ZF-DFE and MMSE equalizers. Rough estimations of the complexity of these algorithms are computed in flops that we define as the number of required complex scalar multiplications in order to decode a symbol vector $\mathbf{s}_{N \times 1}$ transmitted over one OFDM subcarrier through the equivalent channel $\mathbf{H}_{N \times N}$ with $N = MT$. A matrix inversion using the standard Gaussian elimination method needs $N^3 + 2N^2 - N$ flops for ZF decoding followed by N^2 flops for multiplying the inverse matrix and the received vector. Even though the matrix inversion does not need to be realized for each received symbol because the variations of the optical channel are slow, we include this operation in the complexity analysis per received codeword.

The complexity of ZF-DFE is fixed by the QR decomposition and the resolution of the linear system in (2.57). An MMSE equalization contains three matrix multiplications and a matrix inversion. All the complexities are compared in table 5.1 along with the complexity estimates of ML decoders. The four SDM schemes use 4-QAM symbols ($q = 4$). It is worthy to note that the complexities of the ZF-DFE and the ZF decoders are dominated by the same term N^3 and are nearly the same for large MIMO systems.

Table 5.1: Comparison of the complexity of different decoders of $N \times N$ MIMO systems.

	ZF	ZF-DFE	MMSE	ML Sphere decoder
$N \times N$	$N^3 + 3N^2 - N$	$N^3 + \frac{7N^2}{2} + \frac{N}{2}$	$3N^3 + 3N^2 - N$	$\approx N^6$
3-mode (No coding, $N = 3$)	51	60	105	729
TAST 3×3 ($N = 9$)	963	1017	2421	5.31×10^5
6-mode (No coding, $N = 6$)	318	345	750	46656
TAST 6×6 ($N = 36$)	50508	51210	1.44×10^5	2.18×10^9

5.1.1 Performance of ZF decoding

We start by testing ZF decoding of the ST-coded schemes. We run Monte Carlo simulations of the 6×6 SDM system impaired by MDG at amplifiers using the same parameters as in section 4.2.3.

5. LOW-COMPLEXITY SCHEMES FOR LARGE OPTICAL MIMO SYSTEMS

The average BER curves of both, uncoded (NC) and TAST-coded schemes for three fiber coupling strengths are given in Fig 5.1. In all scenarios, mode scramblers are used at FMAs. First, for the uncoded scheme given by plain lines, it is obvious that the ZF decoder (circle-marked) performs worse than the ML decoder (square-marked) because of the enhancement of noise through the channel inversion operation.

On the other hand, when TAST coding is used (dashed curves) along with ZF decoding, a performance gain is obtained in all cases compared to the ZF-decoded NC scheme. At $\text{BER} = 10^{-3}$, the ZF-decoded ST scheme has a penalty of 1.3 dB and performs worse than the optimal ML-decoded NC scheme that has a penalty of 0.6 dB for medium and strong coupling. The similar performance is due to the same MDL statistics obtained in both coupling cases when mode scramblers are used, as demonstrated in the PDFs of the accumulated MDL in Fig. 4.12.

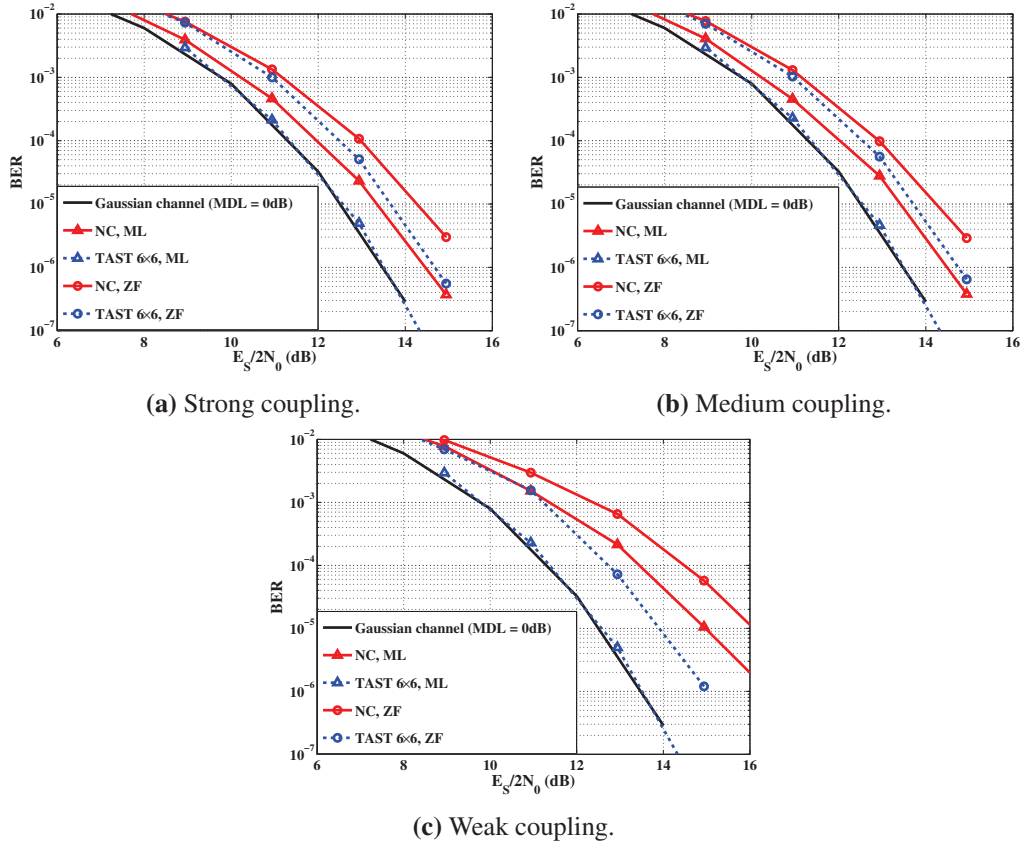


Figure 5.1: Performance in terms of average BER versus SNR of ZF-decoded 6×6 ST schemes with mode scrambling, obtained through Monte Carlo simulations.

However, for weak coupling, ZF-decoded TAST outperforms the NC scheme but a penalty of 1.7 dB compared to the MDL-free case remains. The observed performance of ST codes over the optical SDM system is surprising because it is completely different from the one obtained on a wireless channel where ZF decoding of the codewords would bring absolutely no gain compared

to a ZF decoding of a spatially multiplexed scheme (NC) as we have seen in Fig. 2.11 for the Golden code over the Rayleigh channel. Indeed, a linear ZF decoder does not exploit the diversity of a Rayleigh MIMO channel but only removes the interference from the other transmit antennas by performing the channel inversion. In the case of the PDL-impaired optical PDM channel, we have seen from the theoretical study on PDL mitigation that the behavior of the error probability is Gaussian-like and there are no diversity gains because the scheme has already an infinite diversity. A coding gain is sufficient to mitigate PDL. We conjecture that the same behavior is observed over MDL-impaired SDM channels given that the MDL levels are finite and the total optical power is conserved thanks to periodic optical amplification. Hence, there is no diversity gain to be brought and sub-optimal decoders are able to achieve a part of the coding gains of ST codes.

5.1.2 Performance of ZF-DFE decoding

Second, we run Monte Carlo simulations to obtain the performance of the ZF-DFE decoder on the same 6×6 SDM system impaired by MDG at amplifiers. We remind the reader that the decoding steps of the ZF-DFE were presented in section 2.4.4.2. They consist in performing a QR decomposition of the channel matrix followed by a successive decoding of the components of the symbol vector $\mathbf{s}$ using a new real-valued equivalent channel given by an upper triangular matrix:

$$\begin{aligned} \tilde{\mathbf{y}}_{n \times 1} &= \mathbf{R} \mathbf{s}_{n \times 1} + \tilde{\mathbf{n}}_{n \times 1} \\ \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_n \end{bmatrix} &= \begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ 0 & r_{2,2} & \cdots & r_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{n,n} \end{pmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} + \begin{bmatrix} \tilde{n}_1 \\ \tilde{n}_2 \\ \vdots \\ \tilde{n}_n \end{bmatrix} \end{aligned} \quad (5.1)$$

where $n = 2N = 2MT$. Starting from the last component, each data symbol s_i is decoded and its contribution subtracted from the previous equation before decoding the symbol s_{i-1} . The average BER curves of both, NC and TAST-coded schemes for the strong and weak fiber coupling strengths are given in Fig 5.2. Since the use of mode scramblers gives the same MDL statistics in strong and medium coupling fibers, we omit the results for the latter. Again, it is evident that the ZF-DFE decoder performs worse than the ML decoder for the NC scheme because of noise enhancement while solving the system in Eq. (2.57). However, the successive interference cancellation in the ZF-DFE gives it a performance advantage compared to ZF decoding.

On the other side, when ST coding is used along with ZF-DFE decoding, it achieves the same penalty of 0.4 dB as the optimal ML-decoded NC scheme for medium and strong coupling at $\text{BER} = 10^{-3}$. Furthermore, it outperforms the NC scheme for weak coupling, providing a coding gain of 1.2 dB. This is again surprising compared to the performance over a Rayleigh fading channel for the same reasons mentioned for the ZF decoder. For both coupling strengths, the ZF-

5. LOW-COMPLEXITY SCHEMES FOR LARGE OPTICAL MIMO SYSTEMS

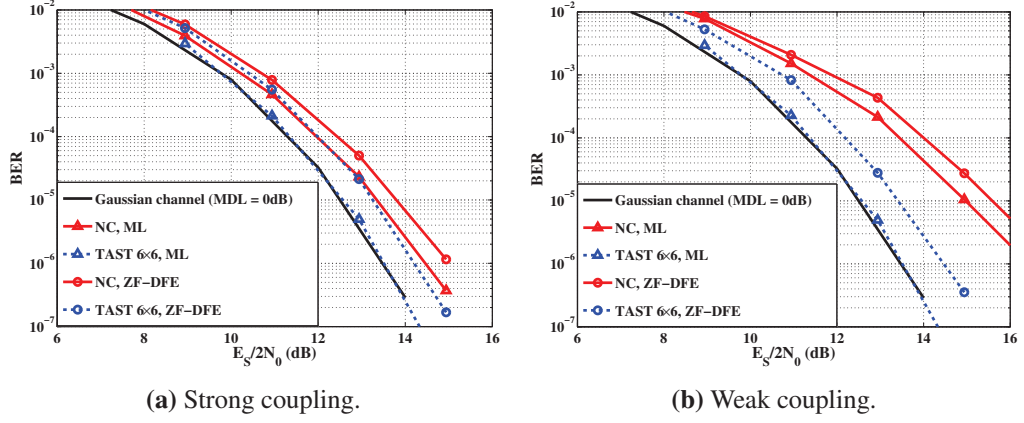


Figure 5.2: Performance in terms of average BER versus SNR of ZF-DFE-decoded 6×6 ST schemes with mode scrambling, obtained through Monte Carlo simulations.

DFE decoder has ~ 0.5 dB advantage over the ZF decoder for an insignificant rise in complexity as shown in table 5.1.

We also observed similar gains for the 3×3 ST-coded and ZF-DFE decoded SDM system. In all investigated scenarios, the low-complexity ZF-DFE decoded ST-coded scheme for a 6×6 SDM system with MDL levels as high as 10 dB is at worst at 1 dB from its corresponding optimally decoded scheme that matches the performance over an MDL-free Gaussian channel.

5.1.3 Performance of MMSE decoding

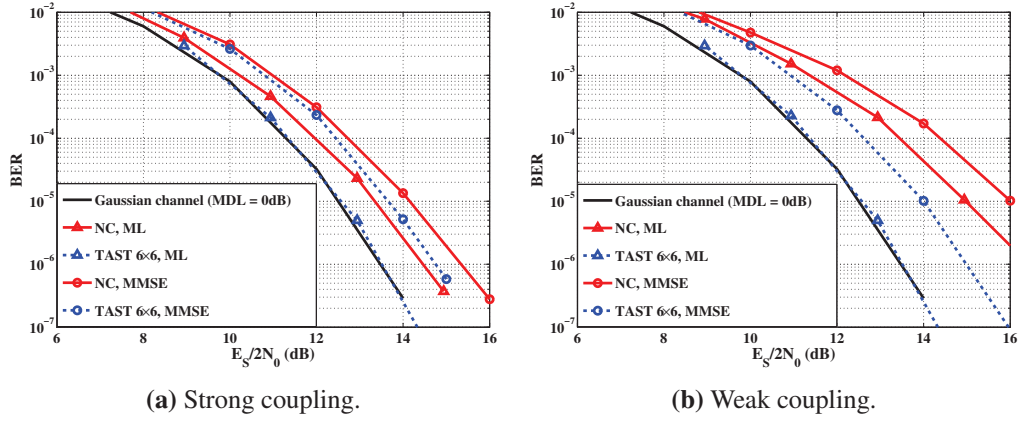


Figure 5.3: Performance in terms of average BER versus SNR of MMSE-decoded 6×6 ST schemes with mode scrambling, obtained through Monte Carlo simulations.

Finally, we test MMSE decoding of the ST-coded 6×6 SDM system with mode scramblers at FMAs. Apart from the channel matrix $\mathbf{H}$, MMSE decoding requires the knowledge of the SNR at the receiver. The average BER curves are given in Fig 5.3. We notice that MMSE yields the same performance as ZF decoding in the uncoded and the TAST-coded scenarios for the investigated

SNR values while requiring three times more operations to perform the equalization. By comparing the linear filters used in the ZF and MMSE decoders in section 2.4.4, we notice that MMSE tries to limit the noise amplification resulting from the channel inversion by adding the inverse SNR term in its filter. While at low SNR, this might give MMSE a performance advantage, at high SNR, this term is negligible and MMSE will have the same performance as ZF which we observe on the obtained BER curves starting from SNR = 6 dB.

5.1.4 Performance of sub-optimal decoders in PDL-impaired PDM systems

While the complexity of ML-decoded 2×2 Silver- and Golden-coded optical systems is still bearable even if an exhaustive search is considered (use $N = 4$ in the formulas in table 5.1 to compute the complexities), the performance of sub-optimal decoders over SDM systems motivated the simulation of their performance over PDL-impaired 2×2 PDM systems. Hereafter, we show the obtained results for the ZF and ZF-DFE decoders in Fig. 5.4.

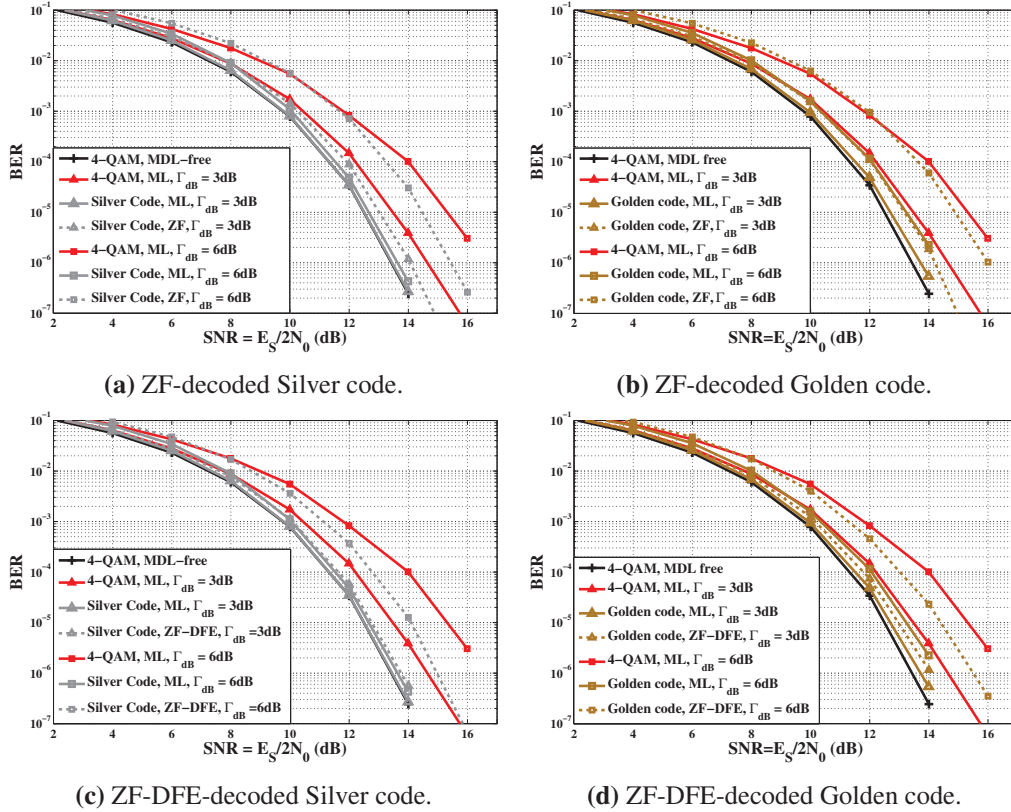


Figure 5.4: Performance in terms of average BER versus SNR of sub-optimal decoding of PDL-impaired 2×2 ST schemes for $\Gamma_{dB} = \{3, 6\}$ dB.

At a PDL of 3 dB, the Silver code shows an SNR penalty of 0.4 and 0.2 dB respectively at $BER = 10^{-3}$ along with ZF and ZF-DFE decoding, while an ML decoding of the code absorbs all the penalty. At a PDL of 6 dB, the ZF-decoded Silver code has a penalty of 2 dB as the one of the

ML-decoded spatial multiplexing (independent 4-QAM symbols over the two polarizations) while a ZF-DFE decoding of the Silver code outperforms it and achieves a penalty of 1.5 dB. The same observation can be done for the Golden-coded system.

5.2 Multi-block coding of SDM schemes

The second strategy suggested for reducing the complexity consists in a multi-block space-time coding approach. Again, we consider the previous 6-mode SDM system where FMAs have a non-zero MDG, to study two new ML-decoded full-rate ST configurations. While we showed that a square $\mathbf{X}_{6 \times 6}$ ST-code absorbs all MDL-induced penalty when associated to mode scrambling (Fig. 4.18), the code requires a high decoding complexity and the processing will become intractable for larger SDM systems. Therefore, we suggest to limit the dimension of the ST-coded system by shortening the length of the coding schemes. Lower complexity full-rate schemes over $T < 6$ time slots such as two $\mathbf{X}_{3 \times 3}$ coded blocks over $T = 3$ time slots (Fig. 5.5(b)) or three $\mathbf{X}_{2 \times 2}$ coded blocks over $T = 2$ time slots (Fig. 5.5(c)) can be considered. If we compute the number of flops required to decode these new schemes by applying the formula for ML sphere decoding in Table 5.1, we find that $18^6 \approx 3.40 \times 10^7$ flops are needed when using two ST blocks over 3 time slots and $12^6 \approx 2.99 \times 10^6$ flops when using three ST blocks over 2 time slots. An important reduction of complexity is thus obtained compared to the single-block 6×6 coded scheme. Indeed, the latter requires decoding codewords of 36 data symbols at the receiver whereas the new schemes contain only 18 and 12 data symbols respectively.

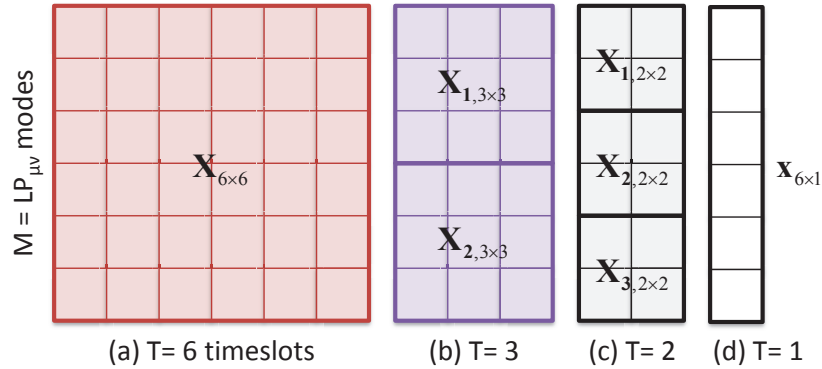


Figure 5.5: Representation of codewords over modes (space) and time for the investigated schemes: (a) Single block with a 6×6 code, (b) Two blocks with 3×3 codes, (c) Three blocks with 2×2 codes, (d) Uncoded spatial multiplexing.

5.2.1 Description of multi-block schemes

To illustrate the benefits of this solution, we define a 2-blocks scheme where each block is coded with the 3×3 TAST code, and a 3-blocks scheme using the 2×2 Silver code. The strategy

for allocating modes to blocks will depend on the modal gain offsets of the FMAs deployed in the optical link. The weakest modes will be coded with a privileged mode in order to average the gain disparities. Hence, for the considered FMA technology with the following gain offsets $\Delta G_{01-11} = -1.3$ dB, $\Delta G_{01-21} = -2$ dB and $\Delta G_{01-02} = -0.2$ dB, the 2-blocks scheme will consist of the LP_{01} and $LP_{21a,b}$ modes in a block, and the LP_{02} with $LP_{11a,b}$ modes in another, whereas the 3-blocks scheme will consist of: $\{LP_{01}, LP_{21a}\}$, $\{LP_{02}, LP_{21b}\}$ and $\{LP_{11a}, LP_{11b}\}$.

For this multi-block coding solution, the vectorized equivalent channel model is the same as the single-block solution in (2.48) with a rearranged equivalent channel matrix $\mathbf{H}_{eq}$ of smaller dimensions due to the use of less time slots $T < 6$. The generator matrix of the coded schemes $\mathbf{M}_G$ depicts the multi-block ST-coded configuration. It will consist of a block diagonal matrix with the generator matrices of the 3×3 TAST code or 2×2 Silver code on its diagonal in the multi-block schemes, as illustrated here:

$$\mathbf{M}_{G,2\text{-blocks}} = \begin{bmatrix} \mathbf{M}_{G,\text{TAST}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{G,\text{TAST}} \end{bmatrix}; \mathbf{M}_{G,3\text{-blocks}} = \begin{bmatrix} \mathbf{M}_{G,\text{Silver}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{G,\text{Silver}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_{G,\text{Silver}} \end{bmatrix} \quad (5.2)$$

From the generator matrices, we can clearly see that these coding schemes would not be optimal over a Rayleigh fading MIMO channel because of their limited transmit diversity gain. Consider the two-blocks scheme with the 3×3 TAST code. The achieved diversity along with ML decoding over a $N_t \times N_r = 6 \times 6$ Rayleigh fading channel will be $d = N_r \times N_t / 2 = 6 \times 3 = 18$ because only 3 copies of each transmitted symbol will be present over 3 different antennas, whereas a 6×6 TAST code can achieve the maximum diversity of the system $d = N_r \times N_t = 36$.

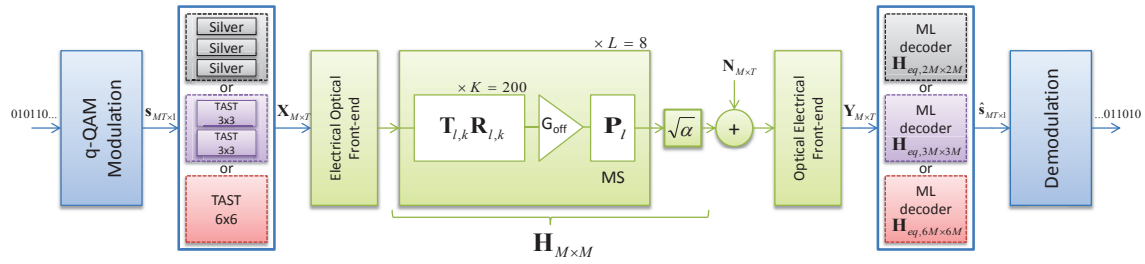


Figure 5.6: Different ST-coded SDM transmission schemes with the simulated optical channel (MS: mode scrambler).

5.2.2 Performance of ML-decoded multi-block schemes

Using Monte Carlo simulations, we evaluate the average BER performance of the ML-decoded multi-block ST schemes over the same previously defined 6×6 SDM system. The transmission system is shown in Fig. 5.6, and Fig. 5.7 shows the BER curves for two different fiber coupling strengths along with mode scramblers after each FMA. The 2-blocks scheme absorbs all the SNR

penalty at $\text{BER} = 10^{-3}$ for strong coupling while the penalty is reduced to 0.5 dB for weak coupling. The 3-blocks scheme shows an SNR penalty of 0.2 dB in strong coupling scenarios and a 1 dB penalty with weak coupling.

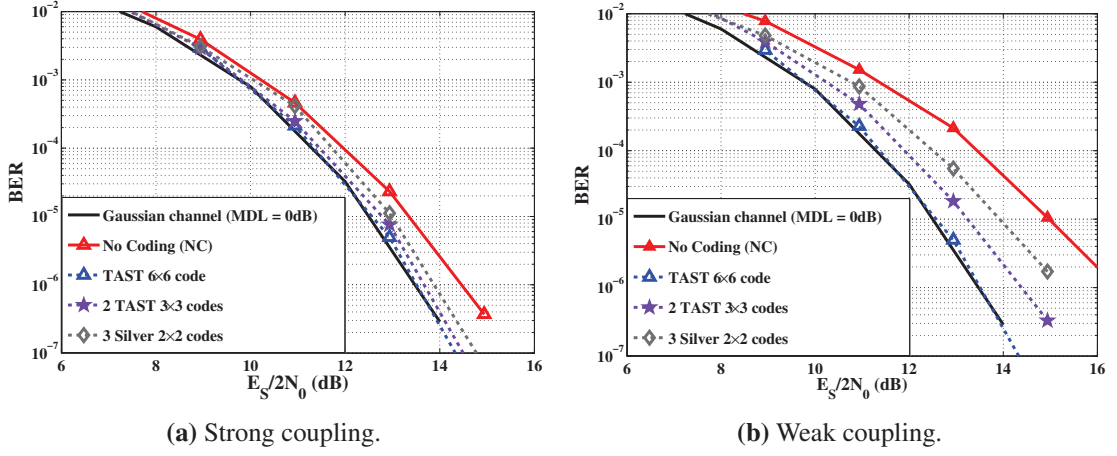


Figure 5.7: Performance in terms of average BER versus SNR of multi-block ST schemes for 6×6 MDM systems with mode scrambling, obtained through Monte Carlo simulations.

The increased penalty of the multi-block schemes is due to the fact that the ST codes average the modal gains inside each block, mitigating intra-block MDL and neglecting inter-block MDL, while the fully coded scheme averages the gains over all the modes. However, the ML decoding complexity is greatly reduced (up to $(2.18 \times 10^9)/(2.99 \times 10^6) \approx 730$ times faster decoding). In conclusion, our second complexity-reducing strategy consisting in multi-block coded schemes reduced the MDL-induced penalty in a 6×6 SDM system, where MDL levels as high as 10 dB are observed, to a maximum of 1 dB from an MDL-free Gaussian channel. While the gains offered by these schemes are lower than the ones of a fully-coded scheme, they are especially impressive due to their complexity and scalability advantages.

If we divide SDM solutions from a crosstalk management point of view as in [176], three main categories can be identified:

- weakly-coupled configurations such as MCFs with distant cores (hence with low spatial density) or some FMFs with low crosstalk (but with high DMGD and scalability issues) where MIMO processing is avoided in order to limit the complexity;
- strongly-coupled configurations such as most FMFs with a DGD minimized by engineering the index profiles and inducing strong coupling that necessitates full-MIMO processing at the receiver;
- and last but not least, strongly-coupled channels divided into weakly-coupled groups such as multi-core fibers with few-mode cores or single-mode MCFs with different inter-core distances.

Among the three categories, the second is shown to achieve the highest spatial density defined as the number of channels packed into a fiber of a given cross section [176]. We have demonstrated, in this section, the benefits of multi-block ST-coding on a 6-mode SDM system belonging to this category. Moreover, we believe that it can be a high-potential solution for both the second and third categories where strongly-coupled configurations can be identified with heterogeneous sources of loss disparities (such as various core-dependent losses/gains...).

5.3 Mode selection for SDM transmission systems

In the previous sections and chapter, we have shown through numerical simulations that ST coding is a promising technique for MDM systems affected by non-unitary crosstalk and mode dependent gains at amplifiers. Both effects result in an accumulated MDL that cannot be compensated with MIMO equalization at the receiver because of the loss of energy and loss of orthogonality between the channels. ST coding, through both optimal and sub-optimal low-complexity schemes, proved its ability to protect the multiplexed information symbols against MDL. However, what if stringent limitations of the complexity of MIMO processing are imposed on transponders due to limited speeds in electronic circuits implementing these algorithms? Would there be any alternative MIMO solutions? Would it be possible to use less modes at the transmitter for multiplexing and less modes at the receiver for signal detection without compromising the system performance?

In [64], Winzer and Foschini analyzed an under-addressed SDM system $\mathbf{H}_{M \times M}$ where $M_T < M$ modes are used at the transmitter and $M_R < M$ at the receiver. In their analysis, strong unitary coupling was assumed and the channel matrix $\mathbf{H}$ was hence unitary. Under these assumptions, they stated that the number of guided modes must match the number of processed modes at the transponders in order to fully exploit the system. This is true for a unitary channel where the capacity of the SDM MIMO system is equal to M times the capacity of a single-mode system. Using less than M modes at the transmitter under-utilizes the multiplexing capability of the channel and detecting less than M modes can result in an important loss of power coupled to the undetected modes. Furthermore, in case the number of addressable modes at the transponders is less than the number of modes supported by the fiber because of limited MIMO processing capabilities, they propose a dynamic switching among modes at the transponders that achieves a better outage probability than selecting a fixed combination (M_T, M_R) of modes.

However, we have seen that imperfections in realistic SDM systems create MDL and lead to a non-unitary SDM channel matrix. MDL reduces the capacity of the channel because the modes are more or less privileged depending on the underlying optical components and fibers used in the SDM link. Hence, it becomes interesting to identify the most privileged modes in a given link configuration and select them for multiplexing at the transmitter and for detection at the receiver.

Furthermore, in case an important portion of the power injected in these modes couples into other propagating modes, initially unexcited at the transmitter, a larger set of modes can be used at the receiver for detection. In [183, 184], a two-mode transmission involving the two degeneracies of the LP_{11} mode propagating at the same velocity is performed on a 3-mode fiber with a large DMGD. The transmission over LP_{01} mode was excluded because of the large DMGD. However, at the receiver, all modes are detected in order to retrieve the energy coupled in the LP_{01} mode due to imperfections in mode multiplexers and de-multiplexers. In [184], the experimental setup consists in a back-to-back configuration (only 7 meters of fiber) and in [183], 20km of large-DMGD low-coupling FMF and mode multiplexers with finite mode-rejection ratios are used.

In multi-antenna wireless communications, antenna selection has proven its worth by providing both a diversity order gain and an SNR gain. Indeed, a judicious choice of fewer transmit antennas can improve the performance and increase the capacity of the system [185, 186]. In this section, we suggest mode selection for MDL-impaired large SDM systems where DMGD is managed optically (low-DMGD fibers) and electronically (OFDM with cyclic prefix + channel shortening techniques). We investigate the system performance when selecting an appropriate set of modes for multiplexing according to an energy criterion. Then, we enhance the system performance by detecting a larger set of modes at the receiver.

5.3.1 Motivation and principle

Mode selection is motivated by the fact that the M propagating modes in an SDM system do not observe the same losses since waveguide imperfections have different impacts on the modes. Furthermore, the optical components inserted in the transmission link may have different modal loss disparities depending on their design. We illustrate this observation through two examples.

First, consider an SDM channel impaired by misalignments at splices and bending losses as in the first case study of the previous chapter. We simulate the channel modeled in (4.4) with $K = 300$ sections of 6-mode fibers with a parabolic index profile, a numerical aperture $NA = 0.205$ and a core radius $r_c = 8.7 \mu\text{m}$. A statistical analysis of the received energies over the 6 modes in 10^6 realizations of the channel matrix in (4.4) gives the probability distribution functions in Fig. 5.8 at two different misalignment levels $\sigma_{x,y}$.

Some of the average received energies per mode in Fig. 5.8 are higher than one because the total received energy is normalized to 6, the total number of modes. One of the selection strategies may consist in defining an energy threshold and choosing the modes having an average received energy higher than the defined threshold. The threshold can be adapted depending on the imposed complexity constraint. For instance, for a misalignment std of $3\%r_c$, the three modes LP_{01} , LP_{11a} and LP_{11b} can be selected and the other modes having the same weak average received energy can be disregarded.

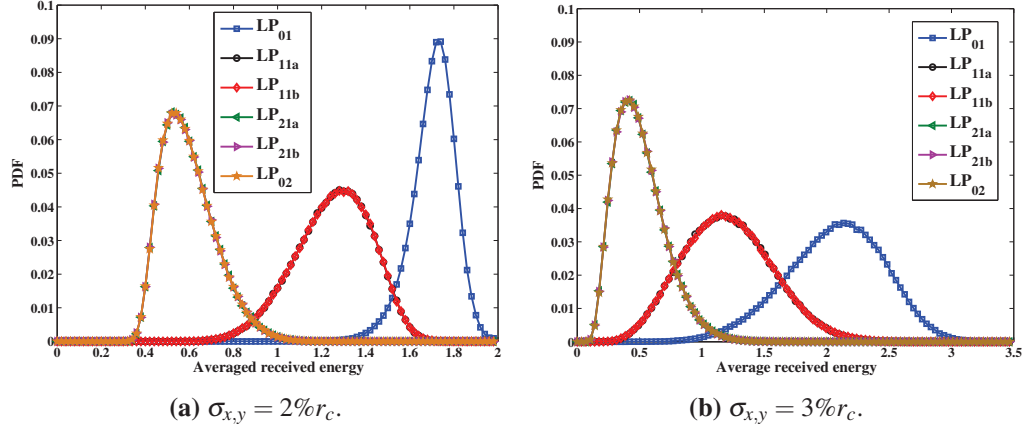


Figure 5.8: PDFs of gains/losses of the 6 modes after propagating in a SDM system made of $K = 300$ misaligned sections with two different misalignment levels.

Second, consider a 6-mode SDM channel impaired by MDG at few-mode amplifiers as in the second case study of the previous chapter. We look into the statistics of the various modal losses/gains and coupling coefficients of the channel matrix in two different crosstalk scenarios. In Fig. 5.9, using a color scale and a 6×6 grid, we show the magnitudes of these coefficients averaged over 10^6 realizations of the optical channel in strong and weak coupling scenarios. The simulation parameters are the same as in the previous chapter.

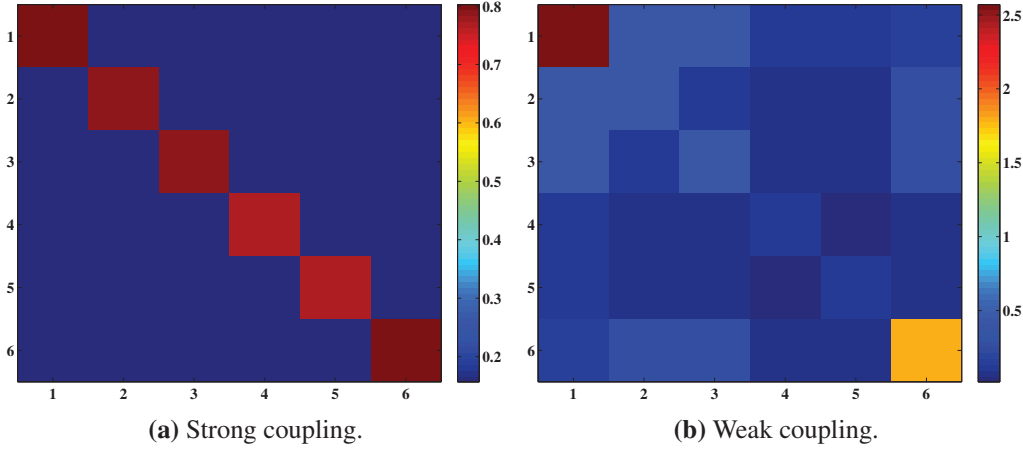


Figure 5.9: Average magnitudes of channel coefficients of a 6-mode SDM system impaired by MDG at few-mode amplifiers in two coupling scenarios (The grid represents the channel matrix $\mathbf{H}$ and $1 = LP_{01}$, $2 = LP_{11a}$, $3 = LP_{11b}$, $4 = LP_{21a}$, $5 = LP_{21b}$, $6 = LP_{02}$).

We notice that in strong coupling scenario, small disparities appear among the modal gains (diagonal of the 6×6 grid representing the channel matrix $\mathbf{H}$) due to the uniform distribution of energy. Hence, mode selection is not very interesting in this case. On the other hand, in weak coupling scenario, the colors show that 4 modes namely LP_{01} , LP_{02} , LP_{11a} and LP_{11b} are more privileged than the others. The grid also shows that all modes do not couple in the same way.

5. LOW-COMPLEXITY SCHEMES FOR LARGE OPTICAL MIMO SYSTEMS

For both SDM channels modeled with an $M \times M$ matrix $\mathbf{H}$, we can select a set of $M_T < M$ modes to multiplex the data symbols at the transmitter and detect $M_R \leq M$ modes at the receiver side as shown in Fig. 5.10. The new MIMO system is given by:

$$\mathbf{y}_{M_R \times 1} = \mathbf{H}_{M_R \times M_T} \mathbf{x}_{M_T \times 1} + \mathbf{n}_{M_R \times 1} \quad (5.3)$$

where the channel matrix $\mathbf{H}_{M_R \times M_T}$ spans a subspace of $\mathbf{H}_{M \times M}$. In wireless environments, transmit antenna selection provides both a gain rate compared to single-antenna systems and a reduced hardware cost because less radio-frequency (RF) circuits would be installed at the transmitter (DACs, electrical amplifiers...). Different algorithms were proposed to find the set of optimal antennas without an exhaustive search over all possible combinations [185, 186]. The same concept is applied to SDM systems with a selection criterion consisting in choosing the M_T modes having the maximal average received energy after propagating in the MDL-impaired channel. At the transmitter, all modes are excited with a unit energy $E_S = 1$ and the average received energy is computed for each mode over random channel realizations. This operation can be done during the design of a given SDM link and implemented once and for all, for instance in prospective long-haul SDM systems. In other applications where a feedback link might be present (which is not the case of current long-distance optical links), mode selection can be realized dynamically in order to adapt to any variation of the MDL statistics in the link.

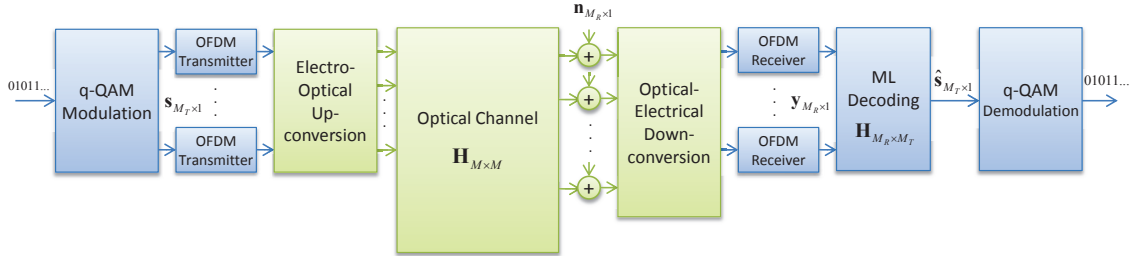


Figure 5.10: SDM transmission system with mode selection.

Mode selection at the transmitter and receiver sides reduces the complexity of MIMO processing at the receiver because fewer symbols need to be decoded simultaneously compared to a transmission over all the modes. In case mode selection is solely applied at the transmitter side and all modes are detected at the receiver side, only a brief increase of complexity is envisaged because we will still be detecting fewer symbols using more observations or more equations in the linear channel model as explained in an application of mode selection to a 10×10 MDM scheme in the next section.

5.3.2 Application to a 10×10 MDM scheme

To show the benefits of mode selection over MDL-impaired optical systems, we consider the MDM channel of the first case study in the previous chapter where MDL is induced by imperfect splices and connections between fiber sections. Three MDM systems are analyzed supporting 3, 6 and 10 propagating modes respectively. $K = 300$ misaligned sections of gradient-index multi-mode fibers with a parabolic profile are concatenated. The simulation parameters are summarized in table 5.2 and the wavelength is fixed to $\lambda = 1550$ nm.

Table 5.2: Parameters of the graded-index fibers used in the simulation.

Number of modes	NA	$r_c(\mu\text{m})$	misalignment std $\sigma_{x,y}(\mu\text{m})$
3	0.205	6	$\{0.26, 0.35\}$
6	0.205	8.7	$\{0.26, 0.35\}$
10	0.205	10	$\{0.26, 0.35\}$

The standard deviations of the misalignments are taken as absolute values because we are comparing fibers with different core radii. They can represent two possible levels of splicing precision. After simulating 10^6 realizations of the 6×6 channel based on the 6-mode fiber and a 10×10 channel based on the 10-mode fiber, the modes with the highest received energies for the 6-mode fiber (resp. the 10-mode fiber) are found to be the LP_{01} , LP_{11a} and LP_{11b} (resp. LP_{01} , LP_{11a} , LP_{11b} , LP_{02} , LP_{21a} and LP_{21b}).

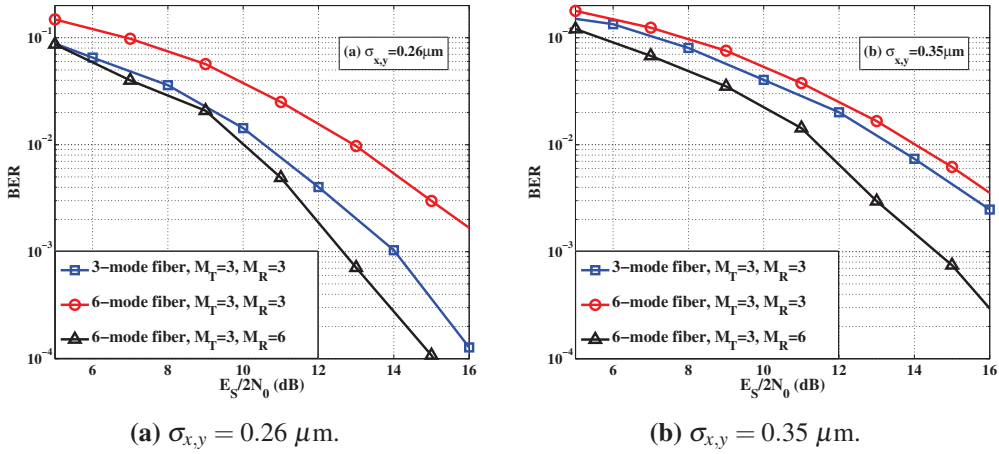


Figure 5.11: Performance in terms of average BER versus SNR of 3-mode fiber and 6-mode fiber using M_T transmit modes and M_R modes at the receiver side, with a spectral efficiency of 6 bits/cu.

We compare the performance in terms of BER curves versus per-mode SNR. The modulated symbols are taken from a 4-QAM constellation. At the receiver, an ML decoder searches for the symbol that minimizes the quadratic distance with the received symbol. In Fig. 5.11, we compare a 3-mode fiber to a 6-mode fiber with the selection of 3 modes ($M_T = 3$) according to

5. LOW-COMPLEXITY SCHEMES FOR LARGE OPTICAL MIMO SYSTEMS

the above-mentioned energy criterion. The same spectral efficiency of 6 bits/cu is achieved in both cases and the average received energy of the 6-mode system is normalized with respect to the 3-mode system. We observe that the 3×3 channel in the 6-mode fiber performs worse than the 3×3 channel on the 3-mode fiber. This is due to important power leakage (coupling) in the other three unexcited modes in the 6-mode fiber as can be seen in the statistics of the received energies per mode in Fig. 5.8b. To retrieve this power loss, we increase the number of detected modes ($M_R = 6$). From the triangle-marked curves, at $\text{BER} = 10^{-2}$, the 3×6 channel has an SNR gain of 0.5 dB (resp. 1.9 dB) compared to the 3×3 channel in the 3-mode fiber for $\sigma_{x,y} = 0.26 \mu\text{m}$ (resp. $\sigma_{x,y} = 0.35 \mu\text{m}$).

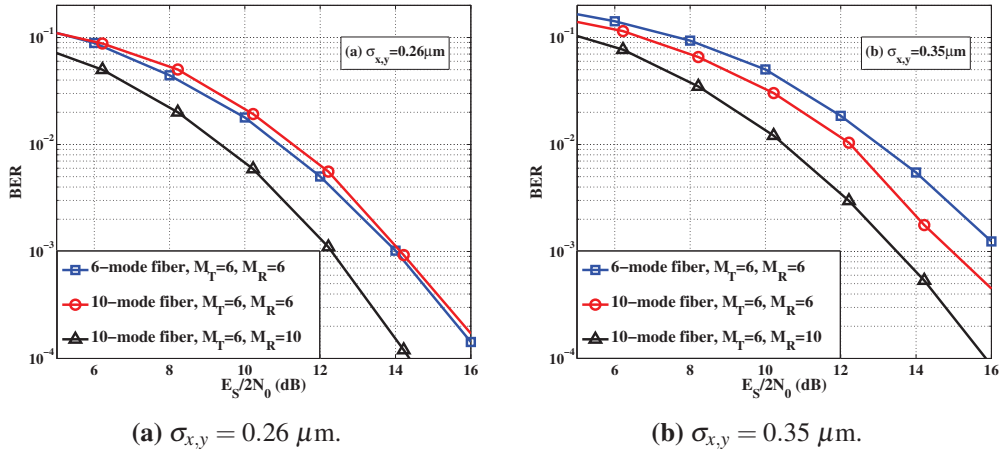


Figure 5.12: Performance in terms of average BER versus SNR of 6-mode fiber and 10-mode fiber using M_T transmit modes and M_R modes at the receiver side, with a spectral efficiency of 12 bits/cu.

In Fig. 5.12, we compare a 6-mode fiber to a 10-mode fiber with the selection of 6 modes ($M_T = 6$), at the same spectral efficiency of 12 bits/cu and the same average received energy normalized with respect to the 6-mode system. At $\text{BER} = 10^{-2}$, we notice that for $\sigma_{x,y} = 0.26 \mu\text{m}$, the performance of the 6×6 channel on the 10-mode fiber is degraded by only 0.4 dB compared to the 6×6 channel based on the 6-mode fiber. On the other side, for $\sigma_{x,y} = 0.35 \mu\text{m}$, the former channel has a gain of 0.8 dB compared to the latter. This gain is explained by a lower power leakage in the unexcited modes that suffer further losses. When we detect all modes in the 10-mode fiber ($M_R = 10$), a gain of 1.5 dB (resp. 2.5 dB) is obtained for $\sigma_{x,y} = 0.26 \mu\text{m}$ (resp. $\sigma_{x,y} = 0.35 \mu\text{m}$) compared to the 6×6 channel on the 6-mode fiber.

The above results show that mode selection is an efficient technique that can be used in SDM systems based on low-DMGD MMFs with a large number of modes. An additional gain can be obtained by increasing the number of detected modes at the receiver side to gain back the power leakage in unexcited modes and therefore improve the performance of the system. While we applied mode selection in an SDM system with misaligned fiber sections, we can foresee a

broader field of applications of this technique. For instance, it is possible to start deploying an SDM system with MMFs supporting a greater number of modes and selecting some modes along with low-complexity MIMO schemes for the moment, while having an infrastructure prepared to welcome higher-complexity MIMO solutions, in the future, in which all modes can be used for multiplexing, once their real-time implementations become possible. Moreover, a combination of mode selection and the previously studied low-complexity ST coding schemes over the selected modes can be envisaged to improve the obtained performance because in this case, ST coding will mitigate the residual MDL between the selected modes.

Summary

In this chapter, we suggested low-complexity MIMO schemes to mitigate MDL in SDM systems. Two strategies were suggested to reduce the decoding complexity of ST-coded systems: on the one hand, low-complexity ZF-DFE decoding was shown to be near-optimal and on the other hand, a multi-block coding approach also showed interesting coding gains and a large potential for scalability. In all investigated scenarios, for MDL levels as high as 10 dB, the low-complexity decoded ST-coded scheme of a 6×6 MDM system is at worst at 1 dB from its corresponding optimally decoded scheme that matches the performance over an MDL-free Gaussian channel. Thus, we traded a portion of the coding gains for a complexity reduction in order to ensure that the DSP remains tractable. Apart from ST coding, we suggested another interesting strategy for large optical MIMO systems impaired by MDL that consists in selecting an appropriate set of modes for multiplexing, when the number of addressed modes at the transceivers is limited by the complexity of MIMO processing in the electronic domain. Although we are not using the full-capacity of the multi-mode fiber when selecting a subset of the propagating modes for multiplexing, we showed through a numerical simulation that, for the same source of imperfection, the obtained performance can match (and sometimes can be better than) the one over a multi-mode fiber with a smaller core supporting the propagation of the same number of selected modes.

Conclusions & Outlook

Conclusions

The growth in bandwidth demand fueled by the increased number of users and machines connected to the Internet will probably motivate major technological changes in future optical transmission systems. Today, polarization multiplexing is a key-enabling technique for 40 Gb/s and 100 Gb/s commercial products whereas a decade ago it was not even on the horizon until dual-polarization coherent detection and de-multiplexing of the polarizations in the electronic domain were demonstrated. This might be also the case for spatial division multiplexing.

In this thesis, we have proposed MIMO coding techniques, known as Space-Time codes, to improve the performance of current and future generations of optical fiber MIMO links. Polarization and space division multiplexing techniques offer both advantageous capacity gains as well as opportunities for the integration of optical components, however at the cost of an increased crosstalk that needs to be digitally managed. Given that this crosstalk is not always unitary and may reduce the capacity of the optical MIMO channels, a simple inversion of the channel matrix is not sufficient to undo the induced penalties.

For 2×2 PDM systems where non-unitary crosstalk is referred to as PDL, we tested the performance of various ST codes, called in this case Polarization-Time codes, taking into account both linear and non-linear effects. The codes were implemented along with an OFDM format to manage dispersion effects and limit the complexity of ML decoding. ST codes performed differently than over a Rayleigh fading channel. The Silver code outperformed the Golden code, the optimal code for 2×2 Rayleigh channels, and the performance of the Alamouti code was insensitive to PDL levels. Moreover, ST codes designed with unitary generator matrices did not show an increased sensitivity to non-linear effects compared to uncoded transmissions.

To perceive the obtained performance, we carried out a theoretical analysis by computing an upper bound of the error probability of a PDL-impaired PDM system. The obtained upper bound showed a Gaussian-like behavior, i.e. the error probability decreased exponentially as a function of the SNR just as the performance over an AWGN channel and unlike Rayleigh fading channels where the error probability decreases as a power of SNR^{-1} , known as the diversity of

the transmission system. Hence, the PDL-impaired PDM system has an infinite diversity for finite levels of PDL and the application of ST codes can bring coding gains by increasing the minimal distance between codewords. Through this analysis, we explained the performance of the investigated codes and found design criteria of optimal PT codes for PDM systems. The minimal distance achieved by the Silver code was found to be larger than that of the Golden code, and the minimal distance of the Alamouti code was found to be independent of PDL. Furthermore, we found that the gains brought by ST codes and FEC codes summed up thanks to the Gaussian-like behavior of the error probability.

After understanding the benefits of PT codes, we conducted a transmission experiment of ST codes over 1000 km of SSMFs emulated using a recirculating loop to take into account more complicated channel effects such as interactions of PDL with PMD and distributed ASE noise. The experimental results validated the numerical simulations and the theoretical analysis proving that the implementation of redundancy-free PT codes in future transmission systems is very beneficial for mitigating PDL levels as high as 6 dB.

Then, we have investigated the use of ST coding in 3-mode and 6-mode SDM systems with various sources of non-unitary crosstalk referred to as MDL. In particular, full-rate TAST codes were able to remove alone the penalties induced by misaligned fiber splices and connections. When high MDL levels up to 10 dB arising from the accumulation of important mode dependent gains at in-line optical amplifiers were observed, TAST codes complemented optical solutions, such as mode scramblers at amplification stages or strong-coupling fibers, in removing the MDL-induced penalties. Hence, ST coding can be implemented as a standalone transmitter-side solution for MDL or as a complementary solution depending on the coupling properties of the installed FMs.

Finally, we were interested in the scale-up possibilities of ST coding for larger SDM systems ($M > 6$). Even though a reduced-complexity lattice decoder, the sphere decoder, was used to implement the ML criterion and optimally decode the ST-coded schemes, the obtained complexity can rapidly increase with the dimensions of the system. Thus, we suggested low-complexity implementations of ST coding that traded a fraction of the optimal performance with a complexity reduction. ZF-DFE decoding and a multi-block coding approach showed both a near-optimal performance with, at worst, a 1 dB SNR gap to their corresponding fully-coded and optimally-decoded schemes. Later, apart from ST coding, we proposed a criterion of transmit-side mode selection over SDM systems impaired by MDL to enhance their performance and limit the decoding complexity and hardware cost of optical transceivers.

Outlook

These observations pave the way for further investigations of the benefits of redundancy-free ST coding in optical MIMO systems. Among the many lines of research, we develop four that we consider as the most interesting:

- Code design for optical MIMO communications: based on the derived criterion for PDM systems, we can search for the coding schemes that best match the orthogonality constraints without reducing the multiplexing gain. Besides, the derived upper bound and design criteria can be further analyzed to include the statistics of PDL such as the Maxwellian distribution of Γ_{dB} . On the other hand, a theoretical analysis of MDL-impaired SDM systems can be carried out to extract design rules of codes dedicated for these systems. To do so, a compact SDM channel model is needed.
- Enhanced SDM channel models and experimental validation: in order to focus on MDL, we disregarded the polarization dimension as well as heterogeneous MCF/MMF SDM solutions. However, we can consider new SDM channel models with different MDL sources and polarization-multiplexed SDM channels where PDL can be added as another performance-limiting effect and test new ST coding schemes. Later on, SDM transmission experiments can be carried out to take into account the interaction of modal dispersion and MDL that will result in a frequency dependent MDL, the non-whiteness of noise for a finite number of noise sources and eventually the non-linear effects for high input power levels.
- Reduced complexity solutions for optical MIMO communications: we have used the sphere decoder as a reduced-search lattice decoder to implement the ML criterion. Other decoders can be tested such as the stack decoder or its low-complexity enhanced spherical-bounded version. A joint decoding of the ST and FEC codes can also be considered. Besides, the multi-block coding approach and mode selection can be extended to hybrid MCF/MMF solutions where various levels of MDL can be observed among subsets of channels depending on the underlying fiber technology. In some of these schemes, a decomposition of the equivalent channel matrix into independent sub-blocks is possible, enabling a low-complexity decoding of the system by processing the sub-blocks in parallel.
- Mitigation of non-linear effects using full-rate MIMO coding: another important research field in the domain of optical transmission systems is the mitigation of non-linear effects. Assessing the benefit of using MIMO coding to achieve that in various optical MIMO schemes (PDM, SDM and WDM) is an interesting research line.

Appendix A

OSNR Sensitivity

We remind the reader of the SNR sensitivities of various modulation formats used in the report, through BER curves averaged over an AWGN channel:

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (\text{A.1})$$

where $\mathbf{n}$ is the noise vector with i.i.d. complex Gaussian components of zero mean and variance $2N_0$ and $\mathbf{x}$ the vector of transmitted symbols with a unit average symbol energy $E_S = 1$.

First, we show the BER curves as a function of the symbol SNR $= E_S/2N_0$ for OOK, BPSK, 4-QAM or QPSK, circular and rectangular 8-QAM, and rectangular 16-QAM.

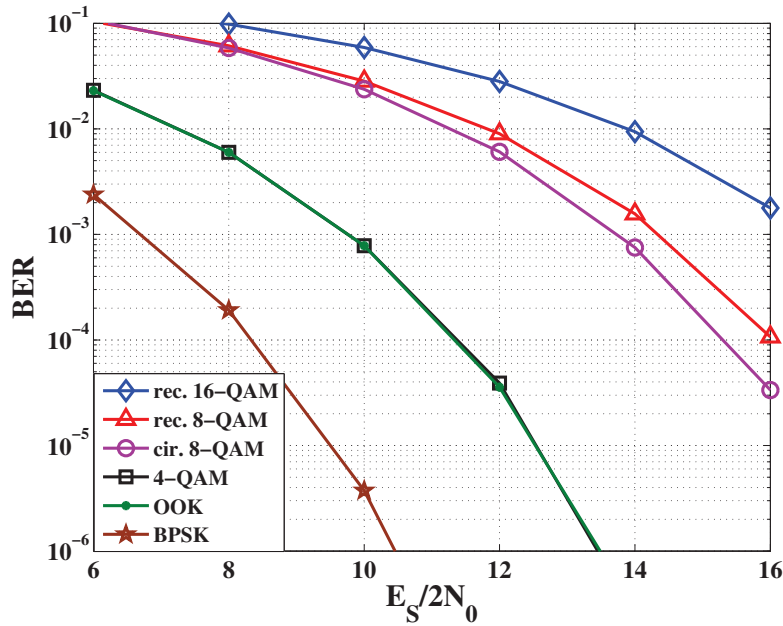


Figure A.1: SNR sensitivities of various modulation formats: average BER versus SNR $= E_S/2N_0$ over an AWGN channel.

While we increase the spectral efficiency by using higher-order modulations, we minimize the minimum distance between constellation points which explains the performance degradation.

The rectangular 16-QAM has a 7 dB penalty compared to 4-QAM. The rectangular 8-QAM has a 4.5 dB penalty compared to 4-QAM while the circular version brings a 0.5 dB enhancement resulting in a 4 dB penalty. As for OOK and BPSK that have the same spectral efficiency, BPSK has a better sensitivity of 3 dB due to its larger minimum distance for the same average energy.

The BER curves can also be given as a function of the per-bit signal-to-noise ratio $\text{SNR}_{\text{bit}} = E_b/2N_0$ where $E_b = E_s/k_c$, k_c being the number of bits carried by each modulated symbol.

Appendix B

Polarization of light

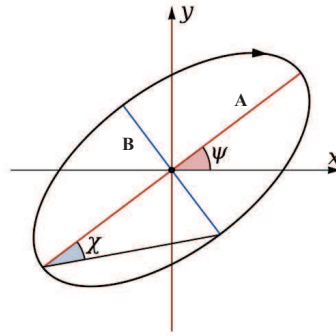


Figure B.1: The polarization ellipse completely describes the polarization state: A is the major semi-axis, B the minor semi-axis, ψ the orientation of the ellipse (tilt or azimuthal angle), $\chi = \arctan(\frac{B}{A})$ the ellipticity angle, and $h = \pm 1$ the sense of rotation (Image source: Wikimedia Commons).

We briefly recall in this appendix the characterization of the polarization of classical electromagnetic waves. The simplest manifestation of polarization to visualize is that of a plane wave, which is a good approximation of most light waves. For plane waves, Maxwell's equations require that the electric and magnetic fields be perpendicular to the direction of propagation and to each other. Conventionally, when considering polarization, the electric field vector is described and the magnetic field is ignored since it is perpendicular to the electric field and proportional to it. The electric field vector of a plane wave may be arbitrarily divided into two perpendicular components labeled x and y (with z indicating the direction of propagation). For a simple harmonic wave, where the amplitude of the electric vector varies in a sinusoidal manner in time, the two components have exactly the same oscillating frequency ω . However, these components have two other defining characteristics that can differ. The two components may not have the same amplitude and may not have the same phase, hence they may not reach their maxima and minima at the same time. In a Cartesian plane, the electrical field $\mathbf{e} = \begin{bmatrix} E_x \\ E_y \end{bmatrix}$ of a sinusoidal wave traveling in a medium can be characterized by the normalized Jones vector:

$$\mathbf{e}_n = \begin{bmatrix} |E_x| \exp(i\phi_x) \\ |E_y| \exp(i\phi_y) \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \exp(i\phi) \end{bmatrix} \quad \text{and} \quad \theta = \arctan \left(\frac{|E_y|}{|E_x|} \right) \quad (\text{B.1})$$

with $\mathbf{e} = \|\mathbf{e}\|\mathbf{e}_n$ and where θ describes the relation between the amplitudes of the electric field components in the x and the y directions and ϕ is the relative phase shift between the electric field components in the x and y directions. Both θ and ϕ define a polarization state of an electric field. At a fixed point in space, the electrical vector traces out an ellipse in the (x,y) plane shown in Fig. B.1. ϕ gives the ellipticity of the state and θ describes the azimuth or the orientation in the (x,y) plane of the major axis of the polarization ellipse. The major and minor axis A and B in Fig. B.1 are related to θ and ϕ by:

$$A = \|\mathbf{e}\| \sqrt{\frac{1 + \sqrt{1 - \sin^2(2\theta)\sin^2\phi}}{2}} \quad \text{and} \quad B = \|\mathbf{e}\| \sqrt{\frac{1 - \sqrt{1 - \sin^2(2\theta)\sin^2\phi}}{2}} \quad (\text{B.2})$$

The orientation of the ellipse is given by $\tan 2\psi = \tan 2\theta \cos \phi$. When $\phi = 0, \pm\pi/2, \pi$, the light is linearly polarized and the ellipse collapses to a line. $\phi = \pm\pi/4, \pm3\pi/4$ defines a circular polarization.

An alternative representation to describe the state of polarization (SOP) of an optical signal is the Stokes parameters $S_{0 \rightarrow 3}$ which can also describe partially polarized or un-polarized light. An additional advantage of using Stokes parameters is that they express the SOP in term of optical powers instead of the electric field, which is more straightforward to measure. The Poincaré sphere is a graphical representation of the SOP that uses the normalized Stokes parameters $S_{1 \rightarrow 3}$ as the axes of a 3-dimensional coordinate system. Fully polarized light can be described by a point on the sphere whereas partially or un-polarized light is described by a point within the sphere. The Stokes parameters are related to the ellipse parameters and the Jones vector representation by the following equations:

$$\begin{aligned} S_0 &= |E_x|^2 + |E_y|^2 = A^2 + B^2 \\ S_1 &= |E_x|^2 - |E_y|^2 = (A^2 - B^2) \cos 2\psi \\ S_2 &= 2\text{Re}(E_x E_y^*) = (A^2 - B^2) \sin 2\psi \\ S_3 &= 2\text{Im}(E_x E_y^*) = 2ABh \end{aligned} \quad (\text{B.3})$$

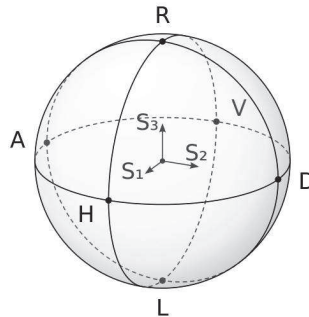


Figure B.2: Representation of the Poincaré sphere with the three Stokes parameter $S_{1 \rightarrow 3}$ (Image source: Wikimedia Commons).

Appendix C

Matrix Operations

The main matrix operations used in the thesis are recalled in this appendix. More details on each decomposition algorithm can be found in [187].

Frobenius norm

The Frobenius norm of an $N \times N$ matrix $\mathbf{H}$ is the Euclidean norm on $\mathbb{C}^N$ defined as:

$$\|\mathbf{H}\|_F = \sqrt{\sum_{i=1}^N \sum_{j=1}^N |h_{ij}|^2} = \sqrt{\text{trace}(\mathbf{H}^\dagger \mathbf{H})} = \sqrt{\sum_{i=1}^N \sigma_i^2} \quad (\text{C.1})$$

where σ_i are the singular values of $\mathbf{H}$ and trace is the trace function (sum of the diagonal elements).

Eigenvalue and Singular Value Decomposition

An eigenvalue and eigenvector of a square $N \times N$ matrix $\mathbf{H}$ are a scalar λ and a non-zero vector $\mathbf{e}_{N \times 1}$ such that:

$$\mathbf{H}\mathbf{e} = \lambda\mathbf{e} \quad (\text{C.2})$$

If $\mathbf{H}$ has N linearly independent eigenvectors $\mathbf{e}_i$ then it can be factorized as:

$$\mathbf{H} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{-1} \quad (\text{C.3})$$

where $\mathbf{E}$ is a square $N \times N$ matrix whose i^{th} column is the eigenvector $\mathbf{e}_i$ of $\mathbf{H}$ and $\mathbf{\Lambda}$ is a diagonal matrix whose elements are the corresponding eigenvalues. If $\mathbf{H}$ is Hermitian ($\mathbf{H}^\dagger = \mathbf{H}$), $\mathbf{E}$ is unitary and $\mathbf{\Lambda}$ has only real values. If $\mathbf{H}$ is unitary, $\mathbf{\Lambda}$ has complex values over the unit circle.

Computing the eigenvalues can be done using the characteristic polynomial:

$$p(\lambda) = \det(\mathbf{H} - \lambda\mathbf{I}) = 0 \quad (\text{C.4})$$

where $\det(\cdot)$ is the determinant of the matrix. Then for each eigenvalue λ_i , the corresponding eigenvector is found by solving:

$$(\mathbf{H} - \lambda_i \mathbf{I}) \mathbf{e}_i = 0 \quad (\text{C.5})$$

using any method for solving a linear system (Gaussian elimination...).

On the other hand, a singular value and pair of singular vectors of a square or rectangular matrix $\mathbf{H}$ are a non-negative scalar σ and two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ such that:

$$\mathbf{H}\mathbf{v} = \sigma\mathbf{u} \quad (\text{C.6})$$

$$\mathbf{H}^\dagger\mathbf{u} = \sigma\mathbf{v} \quad (\text{C.7})$$

There exists a factorization called the singular value decomposition (SVD) of the form:

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger \quad (\text{C.8})$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\mathbf{\Sigma}$ a diagonal matrix with non-negative real numbers. SVD is closely related to the eigenvalue decomposition: the left singular vectors of $\mathbf{H}$ are the eigenvectors of $\mathbf{H}\mathbf{H}^\dagger$, the right singular vectors of $\mathbf{H}$ are eigenvectors of $\mathbf{H}^\dagger\mathbf{H}$, and the non-zero singular values of $\mathbf{H}$ are the square roots of the non-zero eigenvalues of both $\mathbf{H}^\dagger\mathbf{H}$ or $\mathbf{H}\mathbf{H}^\dagger$.

QR decomposition

The QR decomposition also called QR factorization of a real square $N \times N$ matrix $\mathbf{H}$ is a decomposition of $\mathbf{H}$ into an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$ such as:

$$\mathbf{H} = \mathbf{Q}\mathbf{R} \quad (\text{C.9})$$

If $\mathbf{H}$ is non-singular, the factorization is unique. There are several methods for computing the QR decomposition. A popular one is the Gram-Schmidt process. Consider the columns $\mathbf{h}_i, \{i : 1 \rightarrow N\}$ of $\mathbf{H}$. The Gram-Schmidt orthonormalization consists of:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{h}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \\ \mathbf{u}_2 &= \mathbf{h}_2 - \langle \mathbf{h}_2, \mathbf{e}_1 \rangle \mathbf{e}_1, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \\ \mathbf{u}_{k+1} &= \mathbf{h}_{k+1} - \langle \mathbf{h}_{k+1}, \mathbf{e}_1 \rangle \mathbf{e}_1 - \dots - \langle \mathbf{h}_{k+1}, \mathbf{e}_k \rangle \mathbf{e}_k, & \mathbf{e}_{k+1} &= \frac{\mathbf{u}_{k+1}}{\|\mathbf{u}_{k+1}\|}. \end{aligned} \quad (\text{C.10})$$

where $\|\cdot\|$ is the Euclidean norm and $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b}$ is the scalar product of two column vectors. The resulting QR factorization yields:

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N] = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N] \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{e}_1 \rangle & \langle \mathbf{a}_2, \mathbf{e}_1 \rangle & \cdots & \langle \mathbf{a}_N, \mathbf{e}_1 \rangle \\ 0 & \langle \mathbf{a}_2, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{a}_N, \mathbf{e}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle \mathbf{a}_N, \mathbf{e}_N \rangle \end{bmatrix} = \mathbf{Q}\mathbf{R} \quad (\text{C.11})$$

Appendix D

The Sphere Decoder

We consider the lattice $\Lambda = \{\mathbf{H}_{eq,\mathbb{R}} \mathbf{u}_{\mathbb{Z}}\}$ where $\mathbf{u}_{\mathbb{Z}}$ is an $n \times 1$ vector in $\mathbb{Z}^n$, n being the dimension of the lattice. $\mathbf{H}_{eq,\mathbb{R}}$ is the equivalent channel matrix defined in (2.50). The information symbols in the vectorized real-valued transmitted vector $\mathbf{s}_{\mathbb{Z}}$ in (2.50) are taken from a q-QAM constellation. Hence, the search algorithm will look for a point in a subset A of Λ . For simplicity, we drop the indices $\mathbb{R}$ and $\mathbb{Z}$ in the following.

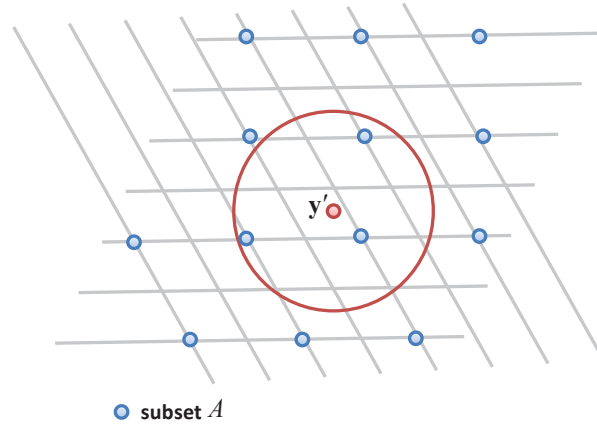


Figure D.1: An illustration of a sphere centered at the received symbol in a 2-dimensional lattice.

The sphere decoder consists in searching for the lattice point in A that is closest to the received point $\mathbf{y}'$, inside an n -sphere of radius C_{SD} centered at the noisy received symbol $\mathbf{y}'$ which is not a lattice point as shown in Fig. D.1. The distance $d(\mathbf{y}', \mathbf{a})$ between any point $\mathbf{a}$ in the subset A and $\mathbf{y}'$ inside the sphere is inferior to C_{SD} .

$$d(\mathbf{y}', \mathbf{a})^2 \leq C_{SD}^2 \quad (\text{D.1})$$

This distance can be expressed as:

$$\begin{aligned}
 d(\mathbf{y}', \mathbf{u})^2 &= \|\mathbf{y}' - \mathbf{a}\|^2 \\
 &= \|\mathbf{y}'^T - \mathbf{a}^T\|^2 \\
 &= \|\mathbf{z}^T \mathbf{H}_{eq}^T - \mathbf{s}^T \mathbf{H}_{eq}^T\|^2 \\
 &= \|(\mathbf{z}^T - \mathbf{s}^T) \mathbf{H}_{eq}^T\|^2 \\
 &= (\mathbf{z}^T - \mathbf{s}^T) \mathbf{H}_{eq}^T \mathbf{H}_{eq} (\mathbf{z} - \mathbf{s})
 \end{aligned} \tag{D.2}$$

where $\mathbf{z} = \mathbf{H}_{eq}^{-1} \mathbf{y}$ is the ZF point. $\mathbf{H}_{eq}$ is a square real matrix that can be rewritten using the QR decomposition into $\mathbf{H}_{eq} = \mathbf{Q}\mathbf{R}$ where $\mathbf{Q}$ is an orthogonal matrix and $\mathbf{R}$ an upper triangular matrix (a Cholesky decomposition can also be used). Now the distance can be further developed:

$$d(\mathbf{y}', \mathbf{a})^2 = (\mathbf{z}^T - \mathbf{s}^T) \mathbf{R}^T \mathbf{R} (\mathbf{z} - \mathbf{s}) \tag{D.3}$$

$$= \|\mathbf{R} (\mathbf{z} - \mathbf{s})\|^2 \tag{D.4}$$

$$= \sum_{i=1}^n \left(R_{ii}(z_i - s_i) + \sum_{j=i+1}^n R_{ij}(z_j - s_j) \right)^2 \tag{D.5}$$

If we note $q_{ii} = R_{ii}$, $i = 1 \dots n$ and $q_{ij} = R_{ij}/R_{ii}$, (D.1) can be expressed as:

$$\sum_{i=1}^n q_{ii} \left((z_i - s_i) + \sum_{j=i+1}^n q_{ij}(z_j - s_j) \right)^2 \leq C_{SD}^2 \tag{D.6}$$

The bounds on a coordinate s_i depend on the values of $s_{j>i}$. Hence, we can bound each coordinate s_i , $i = 1 \dots n$:

$$\left[-\sqrt{\frac{\Theta_i}{q_{ii}}} + \Omega_i \right] \leq s_i \leq \left[\sqrt{\frac{\Theta_i}{q_{ii}}} + \Omega_i \right] \tag{D.7}$$

where Θ_i and Ω_i correspond to:

$$\Omega_i = z_i + \sum_{j=i+1}^n q_{ij}(z_j - s_j) \tag{D.8}$$

$$\Theta_n = C_{SD}^2 \tag{D.9}$$

$$\Theta_{i-1} = \Theta_i - q_{ii} (\Omega_i - s_i)^2 \tag{D.10}$$

we note $\mathbf{b}_i = [b_i^{inf} \ b_i^{sup}]$ the upper and lower bound of s_i .

The sphere decoder searches the lattice points in A satisfying these bounds going through the sphere from its surface towards its center. It starts by the n^{th} coordinate whose value is chosen inside the range determined by the bounds $b_n^{inf} \leq s_n \leq b_n^{sup}$. Apart from being bounded, the chosen

coordinate must belong to the considered q-QAM constellation, i.e. taken from A . This constraint can be easily implemented for rectangular q-QAMs but an extra test procedure is needed for exotic q-QAMs such as circular ones.

Next, the bounds and the value of the $(n-1)^{th}$ coordinate are computed. We choose in the same way all the coordinates s_i until we reach s_1 and obtain the n coordinates of a lattice point $\hat{\mathbf{s}}_1$ inside the sphere. If, at a certain step, a bound inversion occurs, i.e. $b_i^{inf} > b_i^{sup}, i < n$, we have to go up to the step $i-1$ and test other values of s_{i-1} , then find new bounds for s_i . This lattice point search process can be seen as a search in a tree of depth n . Later on, the distance of $\hat{\mathbf{s}}_1$ from the center of the sphere is computed:

$$d(\mathbf{y}', \hat{\mathbf{s}}_1)^2 = C_{SD}^2 - \Theta_1 + q_{11} (\Omega_1 - s_1)^2 \quad (\text{D.11})$$

and the new sphere radius is adjusted to $C_{SD} = d(\mathbf{y}', \hat{\mathbf{s}}_1)$. The search procedure is then repeated in order to find a new lattice point $\hat{\mathbf{s}}_2$ inside the new smaller sphere. At a certain point, no more lattice points will be found inside the sphere, indicated by an inversion of the calculated bounds ($b_n^{inf} > b_n^{sup}$) and the last estimated lattice point is the closest to the received symbol, in other words the ML estimate.

If the initial radius C_{SD} is too small, no lattice points will be found and the radius has to be increased. A judicious choice of the initial radius is given in (2.53) in function of the noise variance and the smallest singular value of the channel. In this case, the decoding complexity of the sphere decoder can be approximated by [156]:

$$\mathcal{O} \left(n^2 \times \left(1 + \frac{n-1}{4dC_{SD}^2} \right)^{4dC_{SD}^2} \right) \quad (\text{D.12})$$

where d^{-1} is a lower bound of the singular values of the matrix $\mathbf{H}_{eq}\mathbf{H}_{eq}^T$. For $C_{SD}^2 \approx d^{-1}$, the decoding complexity can be approximated by $\mathcal{O}(n^6)$. More details on the sphere decoder as well as a flow chart can be found in [155].

Appendix E

Condensed French Version

Techniques Émergentes de Codage Espace-Temps pour les Systèmes de Communications Optiques

E.1 Introduction

Des réseaux d'accès et métropolitains jusqu'aux liens terrestres et transocéaniques de longue distance, les fibres optiques sont les transporteurs inégalés de trafic de données partout dans le monde, principalement en raison de leur capacité à transporter la lumière modulée sur de longues distances avec une faible atténuation et sur une large bande passante. Au cours des vingt dernières années, les systèmes de communication basés sur la fibre optique ont connu une évolution importante et régulière afin de répondre à la demande croissante de débits élevés dans des environnements qui deviennent, de plus en plus, axés sur l'information numérique. Le nombre toujours croissant d'utilisateurs et de machines connectés à Internet affirme l'urgente nécessité d'augmenter la capacité de transport de données de chaque fibre optique installée dans le réseau dorsal (backbone) du système mondial de télécommunications.

Selon le rapport Visual Networking Index (VNI) 2014 de Cisco [4], le trafic IP mondial a quintuplé au cours des 5 dernières années, et va tripler au cours des 5 prochaines années. En particulier, le trafic Internet d'heure chargée (la période de 60 minutes la plus occupée dans une journée) croît plus rapidement que le trafic Internet moyen et atteindra 1 Pétaoctets par seconde (Po/s) en 2018 tandis que le trafic Internet moyen atteindra 311 Téraoctets par seconde. En outre, le trafic dans la région métropolitaine d'une métropole ou d'un pays (à l'échelle de 100 à 1000 km) devrait dépasser le trafic long-haul (couvrant de plus longues distances ≥ 1000 km) en 2015, et représenterait ainsi 62% du trafic total IP vers 2018. Ceci est partiellement due à l'installation extensive des réseaux régionaux de distribution de contenu offrant, entre autres, des services de vidéo à la demande. Les taux de croissance du trafic IP seront également stimulés par le trafic

provenant de nouveaux objets et périphériques connectés tels que les téléviseurs intelligents, les tablettes, les smart-phones et d'autres appareils mobiles ainsi que des services dits "machines à machine".

Afin de répondre à ces demandes croissantes et urgentes de débit, la recherche en communications sur fibre optique a dévoilé un certain nombre de solutions innovantes qui ont conduit au déploiement de systèmes de transmission optique à très haut débit. Ces solutions impliquent l'exploration et l'utilisation de nombreux degrés de liberté dans une fibre mono-mode standard (SSMF) :

1. **Amplitude & Temps** : Les premiers systèmes de transmission optique commercialisés consistaient en une modulation d'intensité (On-Off Keying ou OOK) de la lumière à l'émetteur et une détection directe au niveau du récepteur. Sans aucun traitement avancé au niveau du récepteur, les débits obtenus garantissant une qualité de signal acceptable atteignaient 10 Gigabits par seconde (Gb/s) sur des distances importantes. L'augmentation des débits à 40 et 100 Gb/s en couvrant les mêmes distances de transmission n'était pas possible à cause des exigences de meilleur rapport signal-sur-bruit optique (OSNR) et d'une robustesse diminuée à la dispersion chromatique et la dispersion modale de polarisation (PMD).
2. **Longueur d'onde** : Le développement impressionnant des amplificateurs optiques a permis la transmission de plusieurs longueurs d'onde (ou fréquences) indépendamment modulées sur des milliers de kilomètres sans régénération opto-électrique, un système connu sous le nom de multiplexage en longueur d'onde (WDM) illustré à la fig. E.1a. Les systèmes WDM modernes peuvent transmettre jusqu'à 80-100 canaux en parallèle sur une fenêtre spectrale de 4 THz partitionnée en une grille de 50 GHz. Chaque longueur d'onde peut être modulée à 10, 40 ou actuellement 100 Gb/s, offrant ainsi des capacités totales supérieures à 1 Térabit par seconde (Tb/s) sur une seule fibre.
3. **Phase** : Bien que des modulations binaires de déplacement de phase de type BPSK associées à une détection directe ont été envisagées pour renforcer la sensibilité OSNR par rapport à l'OOK, une solution innovante offrant une capacité accrue et plus de robustesse aux effets linéaires du canal optique était nécessaire. La bien-connue détection cohérente est réapparue comme un outil-clé pour les liaisons optiques à haute capacité, forte des innovations dans les domaines des dispositifs électroniques fonctionnant à haute vitesse et du traitement numérique des signaux. En donnant accès à la fois, à l'amplitude et à la phase du champ optique, les récepteurs cohérents permettent d'augmenter à la fois, le débit et la distance de transmission, des liens optiques grâce à l'égalisation dans le domaine électronique des effets nuisibles du canal optique, principalement les effets de dispersion, à l'aide des techniques de pointe de traitement numérique du signal (DSP). Des débits binaires nets de 40 Gb/s par

longueur d'onde et par fibre ont été obtenus sur des milliers de kilomètres en utilisant une modulation par déplacement de phase en quadrature (QPSK).

4. **Polarisation** : La détection cohérente peut également donner accès à l'état de polarisation du champ électrique de l'onde optique. Tout champ électrique peut être exprimé sous forme vectorielle en fonction de deux états de polarisation orthogonaux arbitraires (appelés aussi modes de polarisation). Le multiplexage en polarisation (PDM ou Pol-Mux) illustré à la fig. E.1b, a permis l'augmentation du débit de transport de données jusqu'à 100 Gb/s par longueur d'onde par fibre. Cette technique consiste à envoyer des symboles de données indépendants sur deux états de polarisation orthogonaux, et donc de doubler l'efficacité spectrale de la transmission optique tout en conservant l'infrastructure existante des précédents systèmes optiques à une seule polarisation. En effet, sur la même largeur de bande spectrale, avec les mêmes sensibilités OSNR et les mêmes taux d'échantillonnage au niveau des transpondeurs (soit l'exemple de circuits intégrés fonctionnant à 28 GSymboles/s ou GBaud), un débit binaire total de $2 \text{ pol} \times 2 \text{ bits/pol} \times 28 = 112 \text{ Gb/s}$ peut être atteint par un système PDM-QPSK.

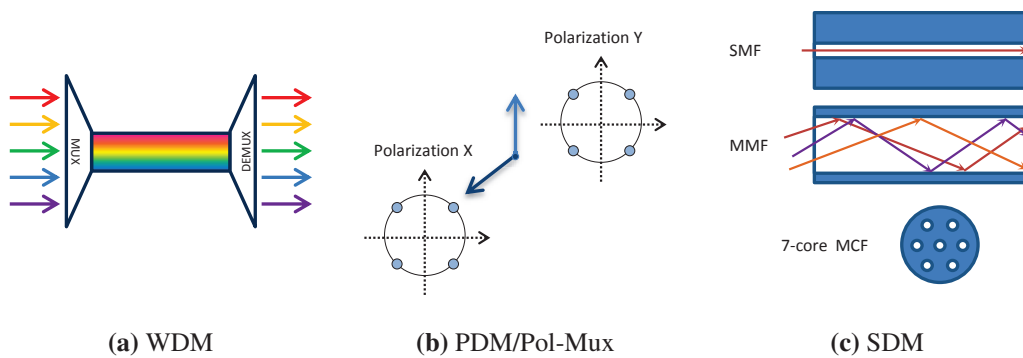


FIGURE E.1: Le multiplexage dans les systèmes de transmission optique : Wavelength Division Multiplexing (WDM), Polarization Division Multiplexing (PDM), Spatial Division Multiplexing (SDM).

Les transpondeurs optiques à 100 Gb/s sont déjà commercialisés, et en R& D, des solutions de faible coût et moins consommatrices d'énergie sont étudiées pour fournir bientôt un débit de 400 Gb/s par longueur d'onde et par fibre sur l'infrastructure actuellement déployée. Des modulations d'amplitude en quadrature à 16 états (16-QAM), un multiplexage par répartition orthogonale en fréquence (OFDM) ou d'autres formats multi-porteuses, ainsi que des techniques avancées de codage correcteur d'erreurs (FEC) sont des candidats prometteurs pour atteindre cet objectif. Cependant, la sensibilité accrue des formats de modulation d'ordre supérieur au bruit et aux nuisances introduites par le canal de transmission nécessitera des niveaux d'OSNR plus élevés, et d'autre part, la mise en œuvre d'algorithmes de décodage de codes FEC plus avancés, requiert des circuits intégrés de plus en plus grands et dissipant plus de puissance. En plus de cela, la capacité des

liaisons optiques à base de fibres mono-mode connaît une limite fondamentale en raison des effets non-linéaires dans la fibre (principalement l'effet Kerr : un changement de l'indice de réfraction de la fibre en fonction de l'intensité du champ électrique se propageant dans celle-ci). Par conséquent, la capacité ne sera pas en mesure de dépasser un certain pic lorsque l'on augmente la puissance du signal injecté dans la fibre ; un pic qui diminue avec la longueur du lien optique [1, 5, 6]. En effet, tandis que les capacités maximales réalisées dans les dernières expériences de transmission double chaque année et demie, le produit distance-capacité maximal double tous les cinq ans [7]. Ainsi, une solution de rupture est nécessaire pour faire face à la demande croissante de capacité.

La dimension **spatiale** est un degré de liberté qui n'a pas encore été utilisé dans la fibre pour le multiplexage de données, et est intensément à l'étude aujourd'hui afin de fournir le prochain gain en capacité. Le multiplexage spatial (SDM) peut être réalisé avec des fibres multi-mode (MMF) présentant un large diamètre de cœur et permettant ainsi la propagation de plus d'un mode spatial à l'opposé des fibres mono-mode actuellement déployées comme illustré à la fig. E.1c. Le multiplexage spatial peut également être réalisé par l'utilisation de fibres multi-cœur (MCF) qui ont plus d'un cœur dans la même gaine. Les modes spatiaux ou cœurs disponibles forment un ensemble orthogonal de canaux sur lequel des données indépendantes peuvent être multiplexées.

Un système de transmission optique WDM/SDM/PDM s'impose fortement comme une solution d'avenir pour atteindre des débits de l'ordre de quelques Pétabits/s dans les futurs systèmes de transmission de données à ultra haut-débit. Outre l'augmentation de la capacité par fibre, un avantage important du déploiement, dans l'avenir, de systèmes SDM plutôt que de l'ajout de plusieurs fibres mono-mode, est la possibilité d'intégrer les composants qui constituent le lien optique (un seul amplificateur optique pour tous les canaux, des transpondeurs intégrés ...). L'intégration permettrait ainsi d'éviter la multiplication du coût et de la consommation d'énergie du système de transmission par le même facteur d'augmentation de la capacité. Cependant, le partage du même espace par les différents canaux amène avec lui de la diaphonie entre ces canaux ainsi que des interactions différentes entre chaque canal et le milieu de propagation, ce qui peut sérieusement dégrader la performance du système. Afin de garantir une communication fiable et cibler de bonnes performances sur ces systèmes à haut-débit densément multiplexés, ces problèmes de diaphonie et de comportements différentiels doivent être résolus.

E.1.1 Objectifs et contributions de la thèse

Dans cette thèse, nous proposons de nouvelles techniques de traitement numérique du signal pour améliorer la performance des solutions de multiplexage pour les systèmes de transmission optique à haut débit d'aujourd'hui et de demain affectés par des effets de diaphonie non-unitaire. Par non-unitaire, nous décrivons les effets provoquant une perte d'orthogonalité et/ou une perte d'énergie entre les canaux sur lesquels les données sont multiplexées. Alors que les transpondeurs cohérents

utilisés actuellement sont équipées d'égaliseurs linéaires qui traitent des effets de dispersion (dispersion chromatique et PMD) qui sont par nature unitaire, ils ne permettent pas de s'affranchir des effets non-unitaires. Plus particulièrement, nous nous intéressons à combattre les pertes dépendant de la polarisation (resp. des modes spatiaux) qui limitent la capacité d'un système PDM (resp. SDM), en utilisant des techniques de codage conçues pour les canaux à entrées multiples et à sorties multiples (MIMO).

En 2010, Mumtaz *et al.* [8] ont montré qu'une famille de codes dits codes Espace-Temps, initialement conçus pour les systèmes multi-antenne sans fils, peuvent efficacement atténuer l'impact de la perte dépendante de la polarisation (dite polarization dependent loss ou PDL) dans les systèmes PDM. Des simulations numériques et une démonstration expérimentale préliminaire ont été menées. Les performances obtenues sur le canal optique ont montré des différences significatives en termes de taux d'erreur binaire TEB (ou bit error rate BER) si on les compare aux performances obtenus sur un canal sans fil à évanouissement de type Rayleigh. Ces résultats ont initié ce travail de thèse qui a commencé en octobre 2011 et dont l'objectif est de développer davantage les résultats obtenus et de proposer des solutions de codage spatio-temporel (ou ST pour Space-Time) pour les systèmes de transmission à multiplexage spatial SDM.

Dans ce projet de thèse, un cadre de simulation a été construit où un système de transmission PDM-OFDM est défini et la résistance des codes ST aux effets linéaires et non-linéaires du canal optique a été évaluée. Le format OFDM a été choisi comme format de modulation préféré vu qu'il assure une mise en œuvre à faible complexité des solutions de codage ST. Ensuite, nous avons utilisé un modèle global de canal optique affectée par la PDL pour calculer une limite supérieure de la probabilité d'erreur de la transmission sur ce canal, dans le but d'expliquer les différences observées par rapport au canal sans-fil et d'en déduire des règles de construction de codes ST appropriées au canal optique avec PDL. En outre, sachant que la plupart des systèmes actuels de transmission optique utilisent des codes correcteurs d'erreurs (forward error correcting ou FEC) pour améliorer la performance ou la portée en ajoutant de la redondance à travers des bits de parité qui offrent la possibilité de corriger des erreurs, nous avons ajouté un module FEC à notre système et avons observé une sommation des gains de codage FEC et codage ST. Cette sommation inattendue a également été expliquée à l'aide de l'analyse de la probabilité d'erreur du canal optique. Ces études ont été complétées par une validation expérimentale de la compensation des pénalités causées par la PDL. À cette fin, un banc expérimental constitué d'une boucle à recirculation comportant un composant optique ayant une PDL non-négligeable a été construit dans les laboratoires de Télécom ParisTech. Nous avons démontré, pour la première fois, la compensation des pénalités causées par la PDL distribuée dans un système optique PDM de 1000 km de longueur, en utilisant des codes ST. L'ensemble du travail réalisé est récapitulé par la suite dans la section "Codage Polarisation-Temps pour les systèmes de transmission optique PDM-OFDM".

En un second temps, après avoir analysé en profondeur la réduction des pénalités de la PDL, nous nous sommes intéressés à combattre les pertes dépendant des modes spatiaux (ou Mode dependent loss, MDL) dans les systèmes de transmission avec multiplexage de modes spatiaux (Mode Division Multiplexing ou MDM), une sous-catégorie des systèmes SDM. La MDL peut sérieusement détériorer la performance des systèmes de transmission MDM. Différentes combinaisons de disparités de pertes entre les modes peuvent survenir dans ces systèmes où un nombre maximum de modes est souhaitée afin d'augmenter la capacité. Les progrès les plus récents dans les technologies de fibres optiques ont conduit à la fabrication de fibres qui soutiennent la propagation d'un petit nombre de modes, habituellement 3 ou 6 modes spatiaux, et qui ont des propriétés intéressantes [9, 10]. Ces fibres sont appelées few-mode fibres (FMF) ou fibres à quelques modes. Des composants optiques supportant quelques modes de propagation tels que des amplificateurs optiques ont été également conçus et optimisés. La mise à l'échelle de ces technologies pour supporter un plus grand nombre de modes est encore un véritable défi. Dans notre travail, nous avons considéré deux principales sources de MDL résultant des imperfections dans les fibres FMF (courbures, épissures et connexions...) ou des disparités de gain dans les amplificateurs optiques few-mode (FMA) périodiquement insérée dans un lien pour compenser l'atténuation de la fibre. Dans les deux cas, nous avons réussi à éliminer complètement les pénalités causées par des niveaux importants de MDL en utilisant des codes ST en tant que solutions autonomes ou en complément à certaines techniques optiques réduisant la MDL dans le canal.

Enfin, nous nous sommes concentrés sur l'étude de la complexité et de la mise à l'échelle des solutions de codage ST. Un inconvénient majeur de ces solutions est l'augmentation sensible de la complexité de leur décodage avec la dimensions du système MIMO en question. En négociant une partie des gains de codage obtenus pour réduire d'une manière significative la complexité de décodage, nous avons réussi à définir deux schémas de codage ST quasi-optimaux qui peuvent être facilement adaptés pour les systèmes MIMO de plus grandes dimensions. Par ailleurs, outre le codage ST, et en tenant compte des particularités de certains effets de diaphonie inter-modale dans les fibres multi-mode et les composants optiques associés, nous avons proposé des techniques de sélection de modes pour les systèmes MDM dans lesquels les meilleurs modes peuvent être sélectionnés pour le multiplexage de données, réduisant ainsi la complexité du DSP au niveau des transpondeurs, tout en garantissant une qualité de transmission améliorée. L'essentiel du travail réalisé sur les systèmes à multiplexage spatial est récapitulé par la suite dans la section "Codage Espace-Temps pour les systèmes de transmission optique SDM-OFDM".

E.2 Codage Polarisation-Temps pour les systèmes de transmission optique PDM-OFDM

L'utilisation optimale des nombreux degrés de liberté dans une fibre optique est une solution clé pour l'augmentation exponentielle des demandes de capacité sur les réseaux de fibres optiques métropolitains et de très longue distance. Au cours des dernières années, la détection cohérente [25] a remplacé les systèmes de détection directe offrant ainsi une sensibilité accrue au niveau du récepteur et permettant l'utilisation de formats de modulation multi-niveaux. En outre, le multiplexage en longueurs d'onde (WDM) a rendu possible la transmission de nombreuses longueurs d'onde indépendantes (80 à 100) sur la même fibre offrant une importante croissance de la capacité des réseaux optiques. De plus, le multiplexage en polarisation dans les systèmes à détection cohérente (PDM) [25, 43] a été adopté pour doubler la capacité de la liaison optique en envoyant deux signaux indépendants sur deux états de polarisation orthogonaux, utilisant ainsi l'amplitude, la phase et la polarisation du signal optique transmis.

Cependant, l'augmentation de l'efficacité spectrale d'une transmission optique, définie comme la quantité d'information envoyée par utilisation du canal, en utilisant le multiplexage se produit au détriment d'une possible réduction de la performance du système de transmission. Ceci est principalement dû à des effets de diaphonie et de différentiel de temps de propagation entre les différents canaux ainsi que des pertes différentielles entre les canaux multiplexés comme notamment la PDL (polarization dependent loss) dans les systèmes PDM [50]. Ces effets peuvent être modélisés en utilisant un formalisme MIMO (Multiple Input Multiple Output) pour les systèmes à entrées et à sorties multiples [103] qui ouvre la voie à l'adaptation de techniques matures de traitement numérique du signal (DSP), initialement conçues pour les transmissions sans-fils multi-antenne. Cela incite aussi à la recherche de nouvelles techniques spécifiques aux schémas MIMO dans les canaux optiques.

Dans cette section, nous nous intéressons aux systèmes multiplexés en polarisation (PDM ou Pol-Mux) modélisés par un système MIMO 2×2 . La mise en œuvre de ces systèmes a été rendue possible grâce au développement de l'électronique fonctionnant à haute vitesse qui a permis l'implémentation d'algorithmes DSP dans le domaine électrique au niveau du récepteur pour récupérer les données transmises. Les dépréciations subies par le signal transmis à très haut débit peuvent être classées en deux catégories : linéaires et non-linéaires. Les principaux effets de canal linéaires sont la dispersion chromatique (CD), la dispersion modale de polarisation (PMD), et la perte dépendant de la polarisation (PDL). Tous ces effets peuvent être modélisés en utilisant des matrices de passage 2×2 (connues aussi sous le nom de matrices de Jones). Quant aux effets non-linéaires, ils résultent de l'interaction entre le champ électrique se propageant dans la fibre et la fibre elle-même. Plus la puissance du champ injecté est importante, plus ces effets deviennent nuisibles.

En un premier temps, nous ne les prenons pas en compte dans les études numériques et théoriques à travers lesquelles nous évaluons les pénalités induites exclusivement par les dépréciations linéaires. Cependant, à travers des simulations et des mesures expérimentales, nous observerons les limitations causées par les effets non-linéaires et déterminerons le point de fonctionnement optimal du système de transmission.

Les effets de dispersion préservent l'énergie totale du signal transmis. L'égalisation de ces effets, CD et PMD, a été étudiée et son efficacité a été prouvée dans un contexte de transmission mono-porteuse [25] ainsi que dans les systèmes multi-porteuses de type OFDM [103]. D'autre part, l'effet de la PDL se présente toujours comme une limitation majeure. Dans les systèmes actuels (à détection directe ou cohérente), l'effet de la PDL n'est pas compensé mais uniquement des marges OSNR sont ajoutées pour absorber les pénalités induites et pour assurer un TEB inférieur à une valeur donnée.

Dans les systèmes de transmission de longue distance, la PDL est introduite par des composants optiques ayant des imperfections (amplis et filtres optiques, routeurs, connecteurs...) et atténuant d'une façon inégale les deux états de polarisation du signal incident, induisant ainsi une perte d'énergie et des fluctuations du rapport signal à bruit optique (OSNR). Récemment [8, 166–168], les codes espace-temps (ST pour Space-Time) ont été suggérés pour compenser les effets de la PDL. Appliqués aux canaux MIMO sans-fils multi-antenne, le codage ST consiste en l'envoi vers le récepteur de plusieurs copies d'un symbole de données sur les différentes antennes et à des créneaux temporels consécutifs, au lieu d'envoyer indépendamment les symboles de données sur les différentes antennes. Ce codage améliore la fiabilité de la transmission car les symboles de données seront moins affectés par les effets d'évanouissement du canal sans-fils et atteindront le récepteur dans un assez bon état pour garantir un décodage fiable [140]. La performance de codes ST, ou plus exactement Polarisation-Temps (PT) dans le contexte du canal optique PDM, ont été numériquement évaluée sur un canal optique contenant de la PDL [8], comme celle du code d'or (Golden code) [151], du code d'argent (Silver code) ou du code d'Alamouti. Une démonstration expérimentale préliminaire a validé ces résultats numériques [167]. Il a été constaté que le code d'argent offre une meilleure performance que le code d'or, contrairement au cas d'un canal sans-fil à évanouissement de type Rayleigh. En outre, la performance du code d'Alamouti s'est révélée être indépendante du taux de PDL dans le lien optique [168]. Dans ce qui suit, nous rappellerons ces résultats numériques récents, et nous fournirons une explication théorique de ces constatations. En outre, nous démontrerons aussi la possibilité de mettre en œuvre les codes Polarisation-Temps dans un système PDM-OFDM et validerons leur potentiel à atténuer les effets de la PDL à travers une expérience de transmission optique de longue distance (1000km de fibres SSMF).

E.2.1 Modèle du système PDM-OFDM codé en polarisation et en temps

Nous commençons par décrire le schéma de transmission général dans lequel les codes Polarisation-Temps seront implémentés. Nous définissons également le modèle canal d'une transmission optique impacté par la PDL qui sera simulé et pris en compte dans l'étude théorique.

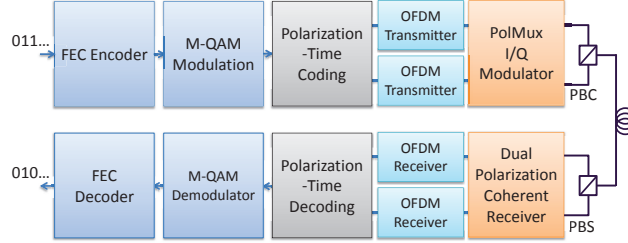


FIGURE E.2: Schéma général d'une transmission PolMux OFDM avec codage Polarisation-Temps.

E.2.1.1 Côté Emetteur

Dans une transmission PolMux OFDM classique [103], deux signaux OFDM indépendamment modulés, où chaque sous-porteuse OFDM transmet un symbole q-QAM, sont envoyés sur deux états de polarisation orthogonaux ; tandis que dans un système codé en polarisation et en temps, les états de polarisation modulés sont corrélés et portent une combinaison linéaire de symboles q-QAM. En effet, le codage Polarisation-Temps consiste à envoyer une combinaison de N_S symboles modulés $s_1 \dots s_{N_S}$ sur deux états de polarisation (Pol_1, Pol_2) au cours de plusieurs créneaux temporels. Par conséquent, si nous considérons, pour simplifier, deux créneaux temporels consécutifs (T_1, T_2), le mot de code peut être représenté par la matrice $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} X_{Pol_1, T_1} & X_{Pol_1, T_2} \\ X_{Pol_2, T_1} & X_{Pol_2, T_2} \end{bmatrix} = \begin{bmatrix} f_1(s_1 \dots s_{N_S}) & f_2(s_1 \dots s_{N_S}) \\ f_3(s_1 \dots s_{N_S}) & f_4(s_1 \dots s_{N_S}) \end{bmatrix} \quad (\text{E.1})$$

Cette opération de combinaison de symboles modulés est représentée par le bloc "Polarization-Time Coding" dans la figure E.2. Les matrices de mots de code des schémas de codage examinés seront données dans la section suivante. La première ligne de $\mathbf{X}$ module le premier état de polarisation et la seconde module le second état orthogonal de polarisation. Les colonnes de $\mathbf{X}$ seront affectées à la même sous-porteuse de deux symboles OFDM consécutifs. Après l'attribution des symboles aux différentes sous-porteuses, un traitement OFDM classique est réalisé comprenant une transformée de Fourier inverse (IFFT) et une insertion d'un préfixe cyclique (CP) bien dimensionné pour absorber tous les délais engendrés par la dispersion (CD et PMD) [103]. Le principal avantage d'utiliser la modulation OFDM pour combattre la dispersion, à la place de solutions mono-porteuse, est la réduction de la complexité de l'étape d'égalisation au niveau du

récepteur au prix d'une légère réduction du débit utile. En effet, l'utilisation d'un CP de longueur appropriée permet de remplacer le filtre égaliseur multi-tap (filtre à réponse impulsionnelle finie multi-étage) dans le domaine temporel habituellement utilisé pour éliminer l'interférence temporelle entre les symboles, par un filtre égaliseur single-tap (à un seul étage) dans le domaine fréquentiel corrigeant les distorsions de phase ou d'amplitude induites par le canal et séparant les deux polarisations pour restituer l'information. Par la suite, chaque signal OFDM est converti dans le domaine optique à l'aide d'un modulateur Mach-Zehnder et un combineur de polarisations combine les deux signaux sur deux états de polarisation orthogonaux du champ optique sur une même fibre.

E.2.1.2 Côté Récepteur

Du côté du récepteur, un séparateur de polarisation sépare le signal optique incident en deux états orthogonaux de polarisation et un récepteur cohérent double-polarisation convertit le signal vers le domaine électrique. Ensuite, un traitement OFDM est effectué, y compris l'enlèvement du CP et une transformée de Fourier (FFT). Chaque sous-porteuse OFDM observe un canal non-dispersif et le symbole reçu peut être exprimé par [103] :

$$\mathbf{Y}_{k,i} = \mathbf{H}_k(\omega_k)\mathbf{X}_{k,i} + \mathbf{N}_{k,i} \quad (\text{E.2})$$

où $\mathbf{X}_{k,i}$ est une matrice 2×2 des symboles transmis sur la k^{me} sous-porteuse ($k = 1 \dots n$) du i^{me} symbole OFDM, $\mathbf{Y}_{k,i}$ est la matrice 2×2 des symboles reçus. $\mathbf{H}_k(\omega_k)$ est la matrice canal 2×2 dans laquelle est inclus le bruit de phase des lasers, la dispersion CD, et les effets de polarisation PMD et PDL. $\mathbf{N}_{k,i}$ est une matrice 2×2 représentant du bruit additif blanc Gaussien. Une séquence d'apprentissage connue par le récepteur est utilisée au début de la trame OFDM pour estimer la matrice de canal par sous-porteuse [43]. Par la suite, l'erreur de phase commune induite par le bruit de phase du laser est corrigée et finalement les symboles de données transmis peuvent être détectés et démodulés. Il est important de noter que les codes spatio-temporels ont été conçus pour fournir des gains de codage et de diversité au système MIMO uniquement lorsqu'ils sont décodés de façon optimale selon le critère de maximum de vraisemblance (ML ou maximum likelihood). Leurs avantages ne sont pas acquis si des décodeurs linéaires sous-optimaux de type ZF (inversion de canal ou Zero Forcing) ou MMSE (minimisant l'erreur quadratique moyenne ou Minimum Mean Square Error) sont utilisés. Considérant le modèle de canal de l'équation. (E.2) et en supposant que tous les mots de code $\mathbf{X}$ sont équiprobables, le décodeur ML offre la performance optimale en estimant le mot de code émis $\mathbf{X}$ par le mot de code $\mathbf{X}'$ selon le critère suivant :

$$\mathbf{X}' = \underset{\mathbf{X} \in \mathcal{C}}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 \quad (\text{E.3})$$

où $\mathcal{C}$ est l'ensemble de tous les mots de code possibles. Les indices k et i ont été omis pour plus de clarté. Le décodage des codes Polarisation-Temps exige que le canal $\mathbf{H}$ reste constant pendant la durée du mot de code (2 créneaux temporels dans notre cas), ce qui est vrai dans le cas d'un système PDM [45].

E.2.1.3 Canal optique

Dans cette sous-section, nous nous intéressons aux propriétés d'un canal optique de longue distance. Nous considérons une liaison optique multi-tronçons avec N_{spans} tronçons, chacun contenant la fibre mono-mode modélisée comme une concaténation d'éléments de PMD ainsi que des éléments de PDL comme on le voit dans la figure E.3. Chaque tronçon est suivi d'un amplificateur à fibre dopée à l'erbium (EDFA), fonctionnant en mode de puissance totale constante, afin d'augmenter la puissance du signal à celle injectée initialement dans le lien au niveau de l'émetteur. Ces amplificateurs ajoutent également du bruit optique provoqué par une émission spontanée amplifiée de photons (ASE). Le bruit ASE accumulé à la fin de la liaison est la source de bruit dominante dans l'ensemble du système de transmission. Dans la suite, le modèle canal de la liaison sera décrit et simplifié afin d'effectuer les études numériques et théoriques de la compensation des effets de la PDL. Nous examinons, en particulier, les statistiques de la PDL, sa dépendance fréquentielle ainsi que les propriétés de bruit vu par le récepteur. Nous ne considérons pas la dispersion chromatique et le bruit de phase des lasers vu que ces deux effets sont communs aux deux états de polarisation.

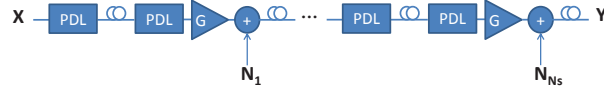


FIGURE E.3: Structure considérée du lien optique de longue distance.

Chaque tronçon de fibre est modélisé par une concaténation de M éléments de PMD aléatoires. La matrice de passage de chaque élément $\mathbf{H}_{PMD}(\omega_k)$ est défini comme le produit d'une matrice diagonale modélisant la biréfringence de l'élément et d'une matrice de rotation [46]. Ces effets unitaires ne provoquent pas de perte d'énergie au signal transmis à la différence de la PDL qui induit des pertes et des fluctuations de l'OSNR. La matrice de passage d'un élément PDL est donnée par :

$$\begin{aligned} \mathbf{H}_{PDL} &= \mathbf{R}_\theta \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\epsilon} \end{bmatrix} \mathbf{R}_\theta^{-1} \\ &= \frac{1}{\sqrt{1+\gamma}} \mathbf{R}_\theta \begin{bmatrix} \sqrt{1+\gamma} & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \mathbf{R}_\theta^{-1} \end{aligned} \quad (\text{E.4})$$

La matrice diagonale donne les valeurs d'atténuation des états de polarisation les moins et plus atténués et $\mathbf{R}_\theta$ est une matrice de rotation aléatoire. L'angle θ est uniformément distribué entre

$[0 : 2\pi]$. ε et γ sont définis par $\Gamma_{dB} = 10\log_{10}\frac{1+\gamma}{1-\gamma} = -10\log_{10}(\varepsilon)$ où $\Gamma_{dB} \geq 0$ est le coefficient de PDL en dB (ou simplement dénommé PDL). Il représente le rapport entre le niveau de perte le plus bas et celui le plus haut. Le facteur de normalisation dans le modèle utilisant la variable γ est communément omis dans la littérature [50, 61] parce que la matrice diagonale et les matrices de rotation sont suffisantes pour tenir compte de l'inégalité des OSNR entre les polarisations ainsi que de la diaphonie.

La valeur individuelle de PDL de chaque composant optique est maintenue aussi faible que possible, c.à.d. $0 \leq \gamma \ll 1$. Cependant, plusieurs composants ayant une PDL non-nulle sont dispersés dans une liaison optique de longue distance, conduisant à une PDL accumulée significative. Un aperçu des travaux antérieurs sur la PDL, en particulier par Mecozzi et al. [53] et Gisin [54], montre des descriptions statistiques de la PDL. Dans [53], il est constaté que la densité de probabilité de la PDL globale Γ_{dB} d'une liaison optique suit une distribution Maxwellienne quand nous considérons de nombreux composants distribués à faible PDL du même ordre. D'autre part, des mesures de la PDL sur des liens déployés [52] ont montré que la distribution de probabilité de Γ_{dB} était Gaussienne et tronquée en raison de la présence d'éléments avec différents niveaux de PDL. Ceux ayant les niveaux de PDL les plus élevés dominaient la statistique de la PDL globale. Dans tous les cas, la PDL globale Γ_{dB} est une variable aléatoire, fonction du nombre d'éléments PDL dans le lien et de leur niveau de PDL.

En second lieu, nous examinons la dépendance fréquentielle de la réponse du canal en raison de la présence de la PMD qui implique une variation en fréquence de la PDL globale. Les fibres optiques actuelles ont un coefficient de PMD PMD_c aussi bas que $0.05ps/\sqrt{km}$. Si nous considérons un lien optique multi-tronçons de longueur $L = 2000km$, la valeur moyenne quadratique de la DGD cumulée (DGD pour Differential Group Dispersion ou différentiel de temps de propagation entre les deux états de polarisation) σ_{DGD} sera égale à $PMD_c\sqrt{L} = 2.2ps$. Cette valeur n'induit pas de fluctuations de la PDL observée par les différentes sous-porteuses d'un signal OFDM occupant une bande fréquentielle inférieure à la bande de cohérence B_{coh} du canal optique : $B_{coh} \propto \sigma_{DGD}^{-1} \approx 450GHz$ [58]. Ainsi, dans la suite, nous considérons un modèle canal plat en fréquence. On notera que la variation fréquentielle de Γ_{dB} en présence de fort niveau de PMD a été observée dans [133, 172]. Dans [133], les auteurs ont constaté que l'insertion de plus de DGD (jusqu'à 100ps de DGD cumulée) dans la liaison optique réduisait les variations des TEB mesurés

fluctuant autour d'une même valeur moyenne de TEB dégradée par la PDL.

$$\begin{aligned} \mathbf{y}_k &= \frac{\mathcal{H}_{1 \rightarrow N_{spans}} \sqrt{2}}{(\text{Tr}(\mathcal{H}_{1 \rightarrow N_{spans}} \mathcal{H}_{1 \rightarrow N_{spans}}^\dagger))^{1/2}} \mathbf{X}_k \\ &+ \sum_{j=1}^{N_{spans}-1} \frac{\mathcal{H}_{j+1 \rightarrow N_{spans}} \sqrt{2}}{(\text{Tr}(\mathcal{H}_{j+1 \rightarrow N_{spans}} \mathcal{H}_{j+1 \rightarrow N_{spans}}^\dagger))^{1/2}} \mathbf{N}_j + \mathbf{N}_{N_{spans}} \\ &= \mathcal{H}_k \mathbf{X}_k + \mathbf{N}_k \end{aligned} \quad (\text{E.5})$$

où $\mathcal{H}_{j \rightarrow N_{spans}} = \mathcal{H}_{N_{spans}} \mathcal{H}_{N_{spans}-1} \dots \mathcal{H}_{j+1} \mathcal{H}_j$ et $\mathcal{H}_j$ étant la matrice de passage du j^{me} tronçon. Chaque bruit $\mathbf{N}_j$ injecté après le j^{me} tronçon est blanc et Gaussien de moyenne nulle et de variance $N_{ASE}/2$ par polarisation et par dimension réelle. A cause de la PDL, $\mathbf{N}_k$, le bruit au récepteur, va être polarisé et sa matrice de covariance $\mathbf{Q} = \mathbb{E}[\mathbf{N}_k \mathbf{N}_k^\dagger]$ ne sera plus proportionnelle à une matrice identité. Pourtant, nous pouvons définir un système équivalent comme [59] :

$$\mathbf{Y}_k = \mathbf{Q}^{-1/2} \mathbf{y}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{N}_k \quad (\text{E.6})$$

où $\mathbf{N}_k$ est un bruit blanc Gaussien de moyenne nulle et de variance $N_0 = N_{spans} N_{ASE}/2$ par dimension réelle. La PDL effective observée par le signal et définie comme le rapport entre les valeurs propres maximale et minimale de $\mathbf{H}_k \mathbf{H}_k^\dagger$ sera plus petite en moyenne que la PDL obtenue à travers le rapports des valeurs propres de $\mathcal{H}_k \mathcal{H}_k^\dagger$.

En appliquant une décomposition en valeurs singulières, $\mathbf{H}_k$ peut être exprimée comme le produit de deux matrices unitaires $\mathbf{U}$ et $\mathbf{V}$ et d'une matrice diagonale avec $\sqrt{1+\gamma_{eq}}$ and $\sqrt{1-\gamma_{eq}}$ comme éléments diagonaux représentant les coefficients de gain/perte des états de polarisations les plus et moins atténués après la propagation dans le lien optique :

$$\mathbf{Y}_k = \mathbf{U} \begin{bmatrix} \sqrt{\lambda_{max}} & 0 \\ 0 & \sqrt{\lambda_{min}} \end{bmatrix} \mathbf{V} \mathbf{X}_k + \mathbf{N}_k = \sqrt{a} \mathbf{U} \begin{bmatrix} \sqrt{1+\gamma_{eq}} & 0 \\ 0 & \sqrt{1-\gamma_{eq}} \end{bmatrix} \mathbf{V} \mathbf{X}_k + \mathbf{N}_k \quad (\text{E.7})$$

où $\gamma_{eq} = (\lambda_{max} - \lambda_{min})/(\lambda_{max} + \lambda_{min})$ et $a = (\lambda_{max} + \lambda_{min})/2$ avec λ_{max} et λ_{min} étant les valeurs propres de $\mathbf{H}_k \mathbf{H}_k^\dagger$. a est une perte commune supplémentaire et inévitable induite par les différents éléments de PDL. Étant intéressé dans notre étude aux pertes différentielles et à la diaphonie engendrées par la PDL, nous mettons de côté l'atténuation commune a . De plus, étant donné que les variations temporelles du canal optique sont plutôt lentes par rapport à la durée d'un symbole OFDM [45], nous ne considérerons que des valeurs constantes de $\Gamma_{dB} = 10 \log_{10} \frac{1+\gamma_{eq}}{1-\gamma_{eq}}$ en première approche pour effectuer les études numériques et théoriques. Ainsi, dans la suite, nous optons pour la matrice $\mathbf{H}$ suivante pour modéliser un canal optique avec de la PDL :

$$\mathbf{H} = \mathbf{R}_\theta \begin{bmatrix} \sqrt{1+\gamma_{eq}} & 0 \\ 0 & \sqrt{1-\gamma_{eq}} \end{bmatrix} \mathbf{R}_\theta^{-1} \quad (\text{E.8})$$

Nous remarquons que nous avons réduit les matrices unitaires $\mathbf{U}$ et $\mathbf{V}$ à de simples matrices de rotation paramétrée par un angle θ . Cela se justifie par le fait que les deux autres paramètres α et β d'une matrice unitaire n'avait aucune influence sur la performance du système (référence : la section 3.2.1 de la version longue du rapport).

E.2.2 Etude numérique de la compensation des effets de PDL

De nombreux codes ST ont été conçus pour les canaux MIMO sans fils. Nous rappellerons les plus célèbres, ceux qui sont les plus performants sur des canaux 2×2 et 2×1 à évanouissement de type Rayleigh : le Golden Code, le Silver code et le code d'Alamouti. Dans la suite, nous donnons la matrice de mots de code $\mathbf{X}$ et le débit de chaque schéma de codage. Le débit du code est défini comme le nombre de symboles transmis par créneau temporel ou time slot (ts). Dans un cas non codé, $\mathbf{X}$ contient simplement 4 symboles q-QAM indépendants et le débit est égal à 2 symboles q-QAM/ts.

E.2.2.1 Le Golden code

Le Golden code [151] offre la meilleure performance sur un canal MIMO 2×2 à évanouissement de type Rayleigh. Ses mots de code sont construits de la façon suivante :

$$\mathbf{X}_g = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha (S_1 + \theta S_2) & \alpha (S_3 + \theta S_4) \\ i\bar{\alpha} (S_3 + \bar{\theta} S_4) & \bar{\alpha} (S_1 + \bar{\theta} S_2) \end{bmatrix} \quad (\text{E.9})$$

où $\theta = \frac{1+\sqrt{5}}{2}$, $\bar{\theta} = \frac{1-\sqrt{5}}{2}$, $\alpha = 1 + i - i\theta$, $\bar{\alpha} = 1 + i - i\bar{\theta}$ et S_1, S_2, S_3, S_4 sont des symboles q-QAM. Le Golden code assure un débit de 2 symboles/ts et n'introduit donc aucune redondance (réduction du débit utile).

E.2.2.2 Le Silver code

Le Silver code offre une performance légèrement inférieure à celle du Golden code pour assurer en contrepartie un décodage plus simple au récepteur. Sa construction se fait de la manière suivante :

$$\mathbf{X}_s = \begin{bmatrix} S_1 + Z_3 & -S_2^* - Z_4^* \\ S_2 - Z_4 & S_1^* - Z_3^* \end{bmatrix}$$

$$\begin{bmatrix} Z_3 \\ Z_4 \end{bmatrix} = \frac{1}{\sqrt{7}} \begin{bmatrix} 1+i & -1+2i \\ 1+2i & 1-i \end{bmatrix} \begin{bmatrix} S_3 \\ S_4 \end{bmatrix} \quad (\text{E.10})$$

où S_1, S_2, S_3, S_4 sont des symboles q-QAM. Le Silver code n'ajoute pas de la redondance et offre à son tour un débit de 2 symboles/ts.

E.2.2.3 Le code d'Alamouti

Le code d'Alamouti offre une performance optimale sur un canal MIMO 2×1 à évanouissement de type Rayleigh. Il peut aussi être utilisé sur les canaux MIMO 2×2 . Le code est défini comme :

$$\mathbf{X}_A = \begin{bmatrix} S_1 & -S_2^* \\ S_2 & S_1^* \end{bmatrix} \quad (\text{E.11})$$

où S_1 and S_2 sont des symboles q-QAM. La matrice de codage a une structure orthogonale qui permet de décoder très facilement ce schéma. Cependant, le code Alamouti introduit une importante redondance sur un canal 2×2 car il n'envoie qu'un seul symbole/ts et réduit le débit de moitié.

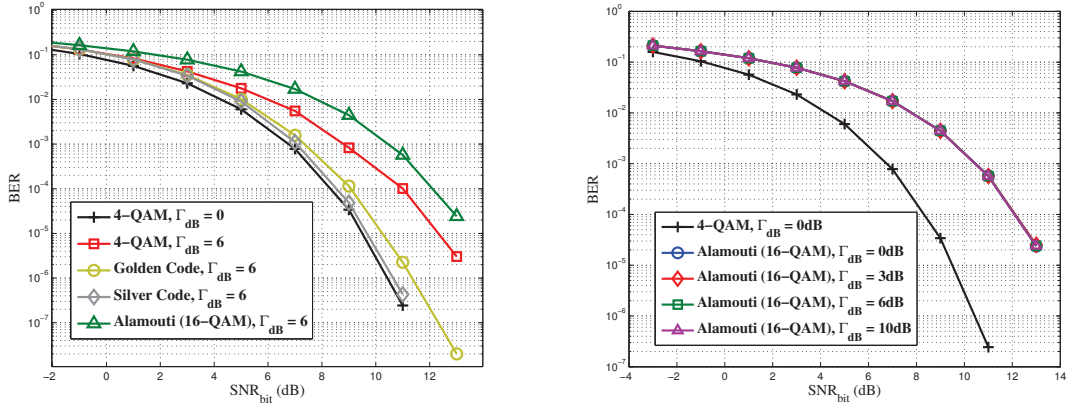
E.2.2.4 Performance des codes sur le canal optique

La performance d'une transmission PDM-OFDM codé en polarisation et en temps sur un canal optique avec de la PDL va être étudiée en utilisant le modèle canal suivant, non-sélectif en fréquence, où $\mathbf{H}$ est définie par l'eq. (E.8) :

$$\mathbf{Y}_k = \mathbf{H}\mathbf{X}_k + \mathbf{N}_k \quad (\text{E.12})$$

Le bruit $\mathbf{N}_k$ est modélisé comme additif et blanc, avec des composantes complexes indépendantes et identiquement distribuées : Gaussiennes et circulairement symétriques $\mathcal{N}(0, 2N_0)$. A l'émetteur, des symboles 4-QAM seront utilisés pour le cas non codé ainsi que pour remplir les matrices de mots de code du Golden code et du Silver code. D'autre part, des symboles 16-QAM, transportant deux fois plus d'information, seront utilisés pour le code d'Alamouti afin de comparer les différents schémas de codage à un même débit binaire de 4 bits d'information/ts. Au niveau du récepteur, un décodage ML est effectué par une recherche exhaustive. Dans le cas d'un régime non codé, il y a 16 combinaisons possibles dans $\mathcal{C}$. Dans le cas du Golden code et du Silver code, $\mathcal{C}$ a $4^4 = 256$ mots de code. Pour le code d'Alamouti, la structure orthogonale nous épargne de la recherche exhaustive et nous pouvons appliquer le critère ML par une seule opération de décodage (simple inversion de canal). Des simulations Monte-Carlo sont effectuées afin d'évaluer la performance des différents schémas de codage en termes de TEB.

La figure E.4(a) montre l'évolution du TEB en fonction du rapport signal sur bruit binaire SNR_{bit} pour une PDL de 6dB. Sans codage MIMO, la dégradation du SNR induit par la PDL est d'environ 2.3dB à un $\text{TEB} = 10^{-3}$. La pénalité SNR - définir comme l'écart par rapport au cas sans PDL pour un TEB donné - est seulement de 0.6dB avec le Golden code et 0.3 dB avec le Silver code, ce qui correspond à un gain de codage de 2dB. D'autre part, la performance du code Alamouti nous montre une pénalité SNR de 3.6dB à un $\text{TEB} = 10^{-3}$. Cela est dû à l'utilisation de symboles 16-QAM ayant une distance minimale plus petite que les symboles 4-QAM à énergie moyenne égale. Sur la figure E.4(b), on observe les performances du code d'Alamouti pour



(a) Performance des régimes :4-QAM, Silver, Gol- den et Alamouti quand $\Gamma_{dB} = 6$. (b) TEB en fonction du SNR_{bit} pour le code d'Ala- mouti à différents Γ_{dB} .

FIGURE E.4: Performance des codes Polarisation-Temps obtenus par simulations Monte-Carlo.

3 niveaux différents de PDL. Nous remarquons que le code offre la même performance indépendamment de la quantité de PDL.

E.2.2.5 Concaténation du codage FEC et du codage MIMO

Alors que les techniques de codage MIMO utilisent les symboles modulés pour former une matrice de mot de code et combattre ainsi certains effets nuisibles du canal, les codes correcteurs d'erreurs, appelés aussi codes FEC (forward error correction) opèrent sur les bits d'information et ajoutent une certaine redondance dans le but de renforcer la performance sur le canal bruité. Le bloc FEC au niveau de l'émetteur précède le bloc de modulation q-QAM, et une unité de décodage FEC correspondant au niveau du récepteur suit le bloc de démodulation. Les codes linéaires en blocs sont une classe importante des codes FEC où une séquence d'information binaire de longueur k est associée (mappée) à une séquence binaire de longueur n appelé un mot de code. Le débit d'un code FEC en blocs r_{FEC} est défini comme :

$$r_{FEC} = \frac{k}{n} < 1 \quad (E.13)$$

La stratégie de décodage optimale des codes FEC est le décodage à décision souple (soft decision decoding ou SDD) [11, Chap.7] qui reçoit à son entrée une valeur réelle non quantifiée indiquant la fiabilité de chaque bit codé (notamment à travers des LLRs pour Log Likelihood Ratios ou Rapports de Vraisemblance Logarithmiques) pour calculer les bits d'information originaux. Une stratégie sous-optimale, mais plus simple à implémenter, est le décodage à décision dure (hard decision decoding ou HDD) [11, Chap.7] qui consiste à quantifier les échantillons à la sortie du démodulateur avant de les envoyer au décodeur FEC.

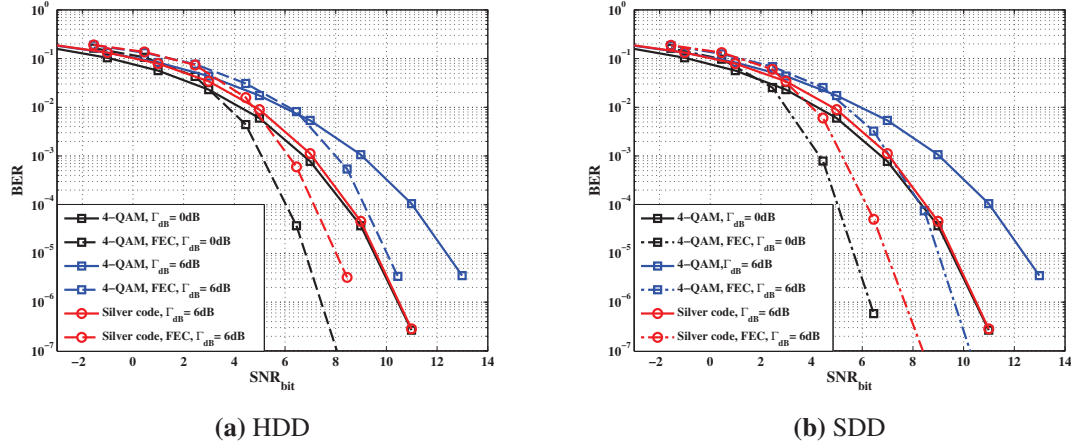


FIGURE E.5: TEB en fonction du SNR_{bit} pour le cas sans codage MIMO et le cas du Silver code, sans et avec codage FEC : (a) Décodage à décision dure HDD), (b) Décodage à décision souple (SDD). Le code FEC simulé est un code BCH(63,45).

Nous nous intéressons dans cette sous-section à l'évaluation du gain total fourni par les codages FEC et MIMO à travers les simulations Monte-Carlo. La performance sur un canal optique avec $\Gamma_{dB} = 6dB$ d'un système 4-QAM non codé et d'un système codé en Silver code est étudiée en leur associant un FEC de la famille de Bose-Chaudhuri-Hocquenghem (BCH) avec $n = 63$, $k = 45$ ayant une capacité de correction d'erreurs $t = 3$. Deux séquences binaires sont indépendamment codées avec le même code BCH, puis attribuées aux modulateurs 4-QAM. Les courbes de TEB obtenues sont reportées sur la figure E.5. Pour un TEB = 10^{-4} , le gain fourni par le Silver code seul est de 2.4dB. Lorsqu'un FEC avec décision dure est considéré, le gain apporté par le code BCH seul est de 1.8dB. La concaténation des deux codes FEC et Silver offrent un gain de codage total de 3.9dB, approximativement égal à la somme des gains de codage séparés $2.4 + 1.8 = 4.2dB$. La sommation des gains de codage FEC et MIMO est également observée lorsque nous utilisons une décision souple. Dans ce cas, le gain offert par le code BCH seul à un TEB = 10^{-4} est de 2.8dB. L'enchaînement avec le Silver code fournit un gain total de 4.9dB et la somme des gains de codage séparés est de $2.4 + 2.8 = 5.2dB$.

Trois résultats majeurs sont retenus à partir des simulations numériques précédentes :

- Le Silver code offre une meilleure performance que le Golden code, contrairement au cas du canal MIMO sans fils.
- La performance du code d'Alamouti est indépendante du taux de PDL dans le lien optique.
- Le gain de codage total obtenu en utilisant un code FEC et un code Polarisation-Temps est égal à la somme des gains apportés séparément par les deux codes.

Ces résultats montrent des comportements différents des codes ST sur le canal optique et une exploration plus approfondie de la compensation de la PDL avec ces codes est nécessaire pour comprendre les résultats obtenus et pour vérifier si les codes étudiés sont optimaux ou si de meilleurs codes peuvent être trouvés pour le canal optique. Pour répondre à cette question, nous menons, dans la sous-section suivante, une analyse de la probabilité d'erreur de la transmission sur le canal optique en calculant une borne supérieure de cette probabilité. Cette borne supérieure constitue un outil approprié pour extraire les critères qu'une matrice de mot de code doit satisfaire afin de minimiser ou atténuer les dégradations du canal de transmission. Un calcul bien connu et largement utilisé de la borne supérieure de la probabilité d'erreur dans le cas d'un canal MIMO à évanouissement de type Rayleigh est présenté dans [140]. Ce calcul avait révélé les critères de construction de codes ST qui ont été utilisés pour construire les Golden et Silver codes.

E.2.2.6 Borne supérieure de la probabilité d'erreur

La probabilité d'erreur est définie comme étant :

$$\begin{aligned} P_{error} &= Pr\{\mathbf{X}' \neq \mathbf{X}\} \\ &= \sum_{\mathbf{X} \in \mathcal{C}} Pr\{\mathbf{X}\} Pr\{\mathbf{X}' \neq \mathbf{X} | \mathbf{X}\} \end{aligned} \quad (\text{E.14})$$

Pour les mots de codes équiprobables, l'application de la borne de l'union nous donne la borne supérieure suivante de la probabilité d'erreur :

$$P_{error} \leq \frac{1}{\text{card}(\mathcal{C})} \sum_{\substack{\mathbf{X}, \mathbf{X}' \in \mathcal{C}, \\ \mathbf{X}' \neq \mathbf{X}}} Pr(\mathbf{X} \rightarrow \mathbf{X}') \quad (\text{E.15})$$

où $\text{card}(\mathcal{C})$ est la cardinalité (taille) de $\mathcal{C}$ et $Pr(\mathbf{X} \rightarrow \mathbf{X}')$ est la probabilité d'erreur par paire (PEP) qui suppose que $\mathbf{X}$ et $\mathbf{X}'$ sont les seuls mots de code possibles dans l'espace des mots de code. Pour calculer la PEP, nous définissons la PEP conditionnelle que nous moyennons sur l'ensemble des états possibles du canal de transmission. Pour un décodage ML, la PEP conditionnelle est définie comme :

$$Pr(\mathbf{X} \rightarrow \mathbf{X}' | \mathbf{H}) = Pr\{\|\mathbf{Y} - \mathbf{H}\mathbf{X}'\|^2 \leq \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 | \mathbf{X}, \mathbf{H}\} \quad (\text{E.16})$$

En utilisant les propriétés du bruit Gaussien N et en appliquant la borne de Chernoff, la PEP peut être majorée par [11, Chap.4] :

$$Pr(\mathbf{X} \rightarrow \mathbf{X}') \leq \mathbb{E}_H \left[\exp\left(-\frac{\|\mathbf{H}(\mathbf{X} - \mathbf{X}')\|^2}{8N_0}\right) \right] \quad (\text{E.17})$$

où $\mathbb{E}_H[\cdot]$ est l'opération de moyennage sur les différentes réalisations possibles du canal.

En moyennant sur $\mathbf{H}$ donnée par l'éq. (E.8) où l'on considère des valeurs constantes de Γ_{dB} et un angle de rotation aléatoire uniformément distribué sur $[0 : 2\pi]$, on obtient :

$$Pr(\mathbf{X} \rightarrow \mathbf{X}') \leq \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) I_0\left(\frac{\gamma_{eq}}{8N_0} \sqrt{a^2 + b^2}\right) \quad (\text{E.18})$$

où $I_0(z)$ est la fonction de Bessel modifiée de première espèce et d'ordre 0. $\mathbf{X}_\Delta = \mathbf{X} - \mathbf{X}' = \begin{pmatrix} \vec{x}_1 \\ \vec{x}_2 \end{pmatrix}$, $\vec{x}_{1,2}$ étant des vecteurs lignes et :

$$a = \|\vec{x}_2\|^2 - \|\vec{x}_1\|^2 \quad (\text{E.19})$$

$$b = 2Re(\langle \vec{x}_1, \vec{x}_2 \rangle) \quad (\text{E.20})$$

Nous pouvons approximer l'expression de la probabilité d'erreur de l'éq. (E.18) pour des fortes valeurs de SNR en utilisant une approximation au premier ordre de $I_0(z)$ quand $z \rightarrow \infty$. Nous obtenons :

$$Pr(\mathbf{X} \rightarrow \mathbf{X}') \approx \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2 - \gamma_{eq} \sqrt{a^2 + b^2}}{8N_0}\right) \quad (\text{E.21})$$

E.2.2.7 Nouveau critère de construction

$I_0(z)$ étant une fonction monotone et croissante sur $z \geq 0$, son minimum correspond à $z = 0$. Cela implique que 'a' et 'b' soient nuls, menant à une probabilité d'erreur indépendante du taux de PDL. En conséquence :

Proposition 1. *Un code Polarisation-Temps peut complètement enlever les effets de la PDL si et seulement si toutes les différences de mots de code satisfont :*

1. $a = \|\vec{x}_2\|^2 - \|\vec{x}_1\|^2 = 0$ et
2. $b = 2Re(\langle \vec{x}_1, \vec{x}_2 \rangle) = 0$.

Quand ce critère de construction est respecté, nous retrouvons la performance de deux canaux AWGN parallèles, ce qui correspond à la meilleure performance réalisable :

$$p_{error,AWGN} \approx \exp\left(-\frac{\|\mathbf{X}_\Delta\|^2}{8N_0}\right) \quad (\text{E.22})$$

Ce critère trouvé est complètement différent des critères du rang et du déterminant minimum établies pour un canal sans-fils à évanouissement de type Rayleigh [140] et définissant respectivement un gain de diversité et un gain de codage sur les courbes de taux d'erreur. Si l'on compare l'approximation de l'expression de probabilité d'erreur du canal optique à SNR élevé de l'équation (E.21), à celle obtenue dans le cas d'un canal de Rayleigh 2×2 , nous constatons que les comportements sont bien différents : la probabilité d'erreur du canal de Rayleigh décroît en SNR^{-2r} [140]

où r est le rang minimum de la matrice $\mathbf{X}_\Delta$. La diversité du système est définie comme la puissance du terme SNR^{-1} ($2r$ dans ce cas) et peut être graphiquement discerné comme la pente des courbes TEB à des valeurs élevées de SNR. En ce qui concerne la probabilité d'erreur du canal optique avec PDL, celle-ci décroît exponentiellement en fonction du SNR comme dans le cas d'un canal à bruit additif blanc Gaussien (éq. (E.22)). Par conséquent, les codes Espace-Temps, utilisés comme codes Polarisation-temps, n'apportent aucun gain de diversité au canal optique avec PDL. Un gain de codage qui sera évaluée par la suite, est seulement introduit, ce qui réduit la pénalité induite par la PDL.

E.2.2.8 Analyse de la performance des codes Polarisation-Temps

La performance des schémas codés et non codés sera examinée en utilisant la borne supérieure de la probabilité d'erreur par paire de l'éq. (E.21). Pour comparer ces schémas, nous calculerons les distances relevées au carré $d^2 = \|\mathbf{X}_\Delta\|^2 - \gamma_{eq} \sqrt{a^2 + b^2}$ de toutes les combinaisons possibles de différences de mots de codes et trouverons, pour chaque schéma, les valeurs minimales de d^2 , que nous appellerons d_{min}^2 . Le meilleur schéma sera celui qui maximise cette valeur minimale d_{min}^2 . Dans le tableau E.1, nous rapportons ces distances minimales calculées pour trois valeurs de PDL. Pour remplir le tableau, nous avons considéré toujours un débit de 4 bits d'information/ts pour tous les schémas et une énergie symbole moyenne $E_s = 1$ pour toutes les modulations utilisées.

TABLE E.1: d_{min}^2 des différents schémas de codage aux différentes valeurs de PDL

	$\Gamma = 0dB$	$\Gamma = 3dB$	$\Gamma = 6dB$	$\Gamma = 10dB$
Silver Code	2	2	2	1.23
Golden Code	2	1.7	1.46	1.07
Non codé	2	1.34	0.8	0.38
Code d'Alamouti	0.8	0.8	0.8	0.8

Tout d'abord, nous constatons que le code d'Alamouti a la même distance minimale pour toutes les valeurs PDL, ce qui explique pourquoi il offre la même performance indépendamment de la PDL (figure E.4(b)). Cela est dû à l'orthogonalité de sa matrice mot de code (Eq. (E.11)) qui donne $a = b = 0$ pour toutes les différences possibles de mots de code. Par conséquent, ce code répond au critère de la *proposition 1*. Toutefois, sa performance est réduite à cause de l'utilisation des symboles 16-QAM pour garantir le même débit utile ; ce qui donne une distance minimale au carré égale à 0.8.

En second lieu, nous notons que le Silver code n'est pas optimale pour combattre la PDL car il ne répond pas au critère de construction dérivé ci-dessus. Contrairement au code d'Alamouti, le Silver code a seulement quelques différences de mots de code ayant $a = b = 0$. Sur la figure E.6, nous traçons la performance du Silver code pour différentes valeurs de PDL. Nous voyons que le

code est capable d'enlever quasiment toutes les pénalités de la PDL tant que le coefficient de PDL est inférieur à 6dB, alors que pour une PDL de 10dB par exemple, le code n'est pas en mesure de pallier toutes les pénalités. Sa d_{min}^2 calculée dans le tableau E.1 explique ce comportement. Pour des PDL de 3 et de 6dB, d_{min}^2 est donnée par la même différence de mots de code ayant $a = 0$ et $b = 0$, et est égale à 2 ; alors que pour une PDL de 10dB, cette distance minimale au carré tombe jusqu'à 1.23 et elle devient dictée par une autre différence de mots de code où $a \neq 0$ et $b \neq 0$.

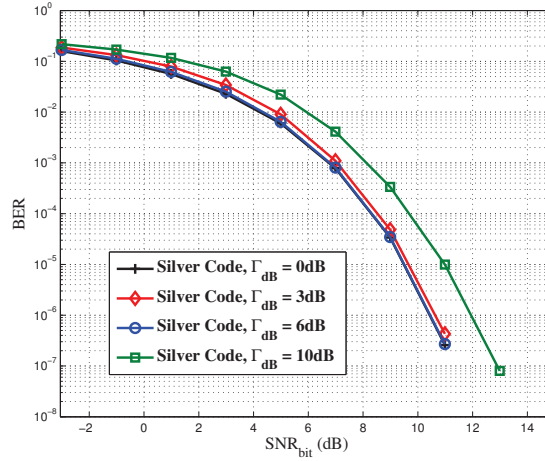


FIGURE E.6: TEB en fonction du SNR_{bit} pour le Silver code, obtenu par simulations Monte-Carlo.

Enfin, sur la figure E.4(a), nous avons vu que le Silver code surpasse le Golden code pour une PDL de 6dB, et que les deux codes l'emportaient sur le cas non codé en réduisant la pénalité causée par la PDL. Encore une fois, ce résultat peut être expliqué par le tableau E.1. En effet, pour un taux de PDL donné, d_{min}^2 est la plus grande pour le Silver code suivie par celle du Golden code, puis par celle du régime non codé fortement réduite par la PDL.

En conclusion, nous étions en mesure d'expliquer, en termes de borne supérieure de probabilité d'erreur, la performance des codes Alamouti, Golden et Silver sur un canal optique avec PDL. Ces codes ont été conçus pour satisfaire aux critères de rang et de déterminant minimal pour un canal sans fils, des critères qui ne sont plus pertinents dans cas du canal optique.

E.2.2.9 Concaténation du codage FEC et du codage MIMO

Les simulations numériques de la performance d'un système Pol-Mux utilisant à la fois un code FEC et un code Polarisation-Temps ont montré que le gain de codage total est égal à la somme des gains de codage séparés apportés par chaque technique. Ce résultat peut être également expliqué théoriquement, et est principalement dû à la décroissance exponentielle de la probabilité d'erreur en fonction du SNR.

La probabilité d'erreur binaire d'un code en bloc linéaire sur un canal AWGN est déterminée par la distance minimale de Hamming du code FEC considéré $d_{min,FEC}$. En partant de l'approximation à fort SNR de la probabilité d'erreur symbole donnée par l'eq. (E.21) et en suivant la démarche exposée dans la section 3.3.2, nous déduisons une borne supérieure de la probabilité d'erreur binaire :

$$P_e \leq K(A_{d_{PT}})^{d_{FEC}} \exp\left(-\frac{d_{FEC}d_{PT}^2 \text{SNR}_{bit} r_{PT} r_{FEC} \log_2 q}{4}\right) \quad (\text{E.23})$$

où $\text{SNR}_{bit} = \frac{E_b}{2N_0}$, E_b étant l'énergie moyenne par bit d'information. K est une constante, $d_{PT}^2 = \underset{\mathbf{X}, \mathbf{X}' \in \mathcal{C}}{\text{argmin}}(\|\mathbf{X}_\Delta\|^2 - \gamma_{eq} \sqrt{a^2 + b^2})$ et $A_{d_{PT}}$ étant le nombre moyen de mots de code situés à une distance d_{PT} d'un mot de code donné. $A_{d_{PT}}$ est souvent appelé kissing number. Selon la stratégie de décodage, d_{FEC} est remplacé par $d_{FEC,HDD} = \lceil \frac{d_{min,FEC}}{2} \rceil$ ou $d_{FEC,SDD} = d_{min,FEC}$. Enfin, r_{PT} est le débit par polarisation et par temps symbole du code Polarisation-Temps égal à 1 pour un rendement plein (full-rate) et 0.5 pour un code à rendement réduit à moitié.

Afin d'évaluer les gains asymptotiques fournis par la concaténation d'un code FEC et d'un code Polarisation-Temps, nous comparons les deux schémas suivants : un premier schéma NC sans FEC ($d_{FEC} = 1$ et $r_{FEC} = 1$) et sans codage MIMO (symboles q-QAM indépendants), et un second schéma $FEC + PT$ utilisant un code en bloc linéaire et un code MIMO à rendement plein. Pour un même probabilité d'erreur donnée, le gain de codage G du schéma codé est égal à :

$$G = \frac{\text{SNR}_{bit,FEC+PT}}{\text{SNR}_{bit,NC}} = \frac{R_c d_{FEC} d_{PT}^2}{d_{NC}^2} \quad (\text{E.24})$$

En décibels, nous obtenons G_{dB} :

$$\begin{aligned} G_{dB} &= 10 \log_{10}(R_c d_{FEC}) + 10 \log_{10}\left(\frac{d_{PT}^2}{d_{NC}^2}\right) \\ &= G_{dB,FEC} + G_{dB,PT} \end{aligned} \quad (\text{E.25})$$

Le premier terme représente le gain de codage du code FEC et le second terme désigne le gain de codage fourni par le codage Polarisation-Temps. L'équation (E.25) montre que le gain asymptotique total obtenu en concaténant un code FEC et un code Polarisation-Temps est égal à la somme des gains fournis par chaque code séparément, validant ainsi le troisième résultat de notre étude numérique.

E.2.3 Validation Expérimentale de la compensation des effets de PDL

Les validations numériques et théoriques de la compensation des pénalités de la PDL en utilisant le codage Polarisation-Temps sont limitées à un canal linéaire avec un seul élément de PDL et du bruit gaussien ajouté au niveau du récepteur. Des résultats expérimentaux initiaux limités au

régime linéaire ont également montré l'efficacité de ces codes à combattre la PDL [167]. Ces simplifications nous ont permis d'analyser le potentiel des codes MIMO pour les systèmes optiques Pol-Mux en utilisant un modèle de canal général et bien défini. L'étape restante consiste à tester la capacité de ces codes à atténuer les effets de la PDL distribuée dans un lien optique à travers une expérience de transmission qui tient compte des interactions de la PDL avec le bruit ASE distribué [60] ainsi qu'en présence d'effets non-linéaires [61].

E.2.3.1 Banc expérimental

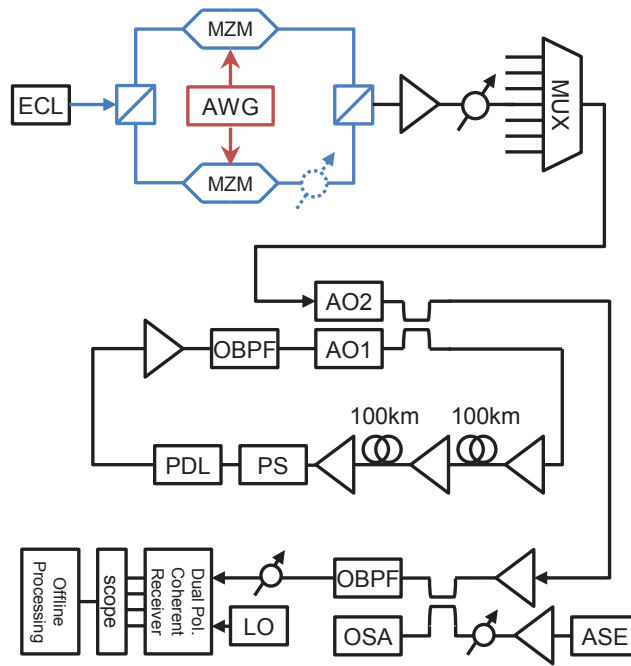


FIGURE E.7: Schéma expérimental. (ECL : External Cavity Laser, AWG : Arbitrary Waveform Generator, MUX : Multiplexeur, AO : Acousto-Optical Modulator, PS : Polarization Scrambler, OBP : Optical Band-Pass Filter, ASE : Accumulated Spontaneous Emission source, LO : Local Oscillator, OSA : Optical Spectrum Analyzer, OSC : Tektronix 50GS/s Oscilloscope).

Le schéma de transmission, représenté sur la figure E.7, se compose d'un émetteur PolMux OFDM, d'une liaison optique de 1000 km et d'un récepteur cohérent double-polarisation. Les signaux OFDM sont formés de 256 sous-porteuses dont 194 portant des données et 10 pilotes estimant l'erreur de phase commune induite principalement par le bruit de phase des lasers. Les sous-porteuses de données sont alternativement modulées avec des symboles 4-QAM, des mots de code Silver, Golden ainsi que du 16-QAM Alamouti. Un préfixe cyclique de longueur 18 échantillons est ajouté à chaque symbole OFDM pour absorber la dispersion. Une séquence d'apprentissage est périodiquement insérée dans la trame OFDM pour estimer le canal et assurer la synchronisation temporelle et fréquentielle [132]. Le débit effectif total de la transmission est de 25Gbits/s alors

que le débit brut est de 36 Gbits/s en tenant compte des différents sur-coûts (CP, séquence d'apprentissage, pilotes). Les signaux OFDM sont générés par un générateur de forme d'onde arbitraire (AWG) ayant deux sorties électriques. Par conséquent, pour pouvoir générer du PolMux corrélé en polarisation, nous nous limitons aux signaux OFDM réels satisfaisant la propriété de symétrie hermitienne. Ensuite, la conversion dans le domaine optique est réalisée en utilisant deux modulateurs Mach-Zehnder (MZM) et les deux signaux optiques sont combinés par un combineur de polarisation, et multiplexés dans le canal optique entre sept autres longueurs d'onde spectralement espacées de 50 GHz.

La ligne de transmission consiste en une boucle à recirculation qui contient deux tronçons de 100km de fibres SSMF, un émulateur de PDL réglé à 2dB de PDL, un brouilleur de polarisation pour randomiser l'état de polarisation entre deux tours successifs dans la boucle et un filtre passe-bande optique. Après l'ajout de bruit ASE supplémentaire au niveau du récepteur, la longueur d'onde portant le signal OFDM est filtrée et le signal est détecté par un récepteur cohérent double-polarisation. Le signal électrique reçu est ensuite acquis avec un oscilloscope temps-réel Tektronix et un traitement off-line sous MATLAB est réalisé pour décoder le signal : un décodage ML et une démodulation suivis par une mesure du taux d'erreur TEB traduit en facteur de qualité Q couramment utilisé dans le domaine des communications optiques :

$$TEB = 0.5\text{erfc}(Q/\sqrt{2}) \quad \text{and} \quad Q_{dB} = 20\log_{10}(Q) \quad (\text{E.26})$$

E.2.3.2 Codage Polarisation-Temps en régime de propagation non-linéaire

Nous transmettons le signal PolMux OFDM codé dans la boucle à recirculation de la figure E.7 et remplaçons l'élément de PDL en ligne par un atténuateur ayant la même perte d'insertion pour émuler, en un premier temps, la PDL localisée au niveau de l'émetteur. Nous insérons pour cela un atténuateur optique variable dans une branche de l'émetteur PolMux (ce qui correspond au pire cas de dégradation de l'OSNR d'un état de polarisation). Tout d'abord, les mesures de TEB sont effectuées dans une configuration back-to-back pour valider la performance précédemment simulée de tous les schémas de codage. Ensuite, les BERs sont mesurés après une transmission à travers 1000km de fibres à des puissances d'entrée variant de -15 à 5dBm. Les courbes de TEB obtenues sont tracées sur la figure E.8 pour le cas 4-QAM non codé et pour le Silver code aux niveaux de PDL de 0, 3 et 6dB.

En premier, nous constatons que toutes les courbes ont leur point de fonctionnement optimal à une puissance d'entrée de -3dBm. Cependant, le régime non-codé souffre clairement de la PDL ($Q = 10.5\text{dB}$ à -3dBm pour une PDL de 6dB) tandis que le Silver code compense totalement la pénalité à une PDL de 3dB, et presque toute la pénalité à une PDL de 6dB ($Q = 12.3\text{dB}$ à -3dBm). En second lieu, les courbes peuvent être divisées en trois régimes : un régime linéaire, jusqu'à -6dBm,

où les effets non-linéaires sont négligeables et les gains de codage correspondent aux résultats numériques obtenus avec un modèle de canal linéaire ; un régime non-linéaire modéré, entre -6dBm et 0dBm, où un TEB minimum est atteint ; et un régime fortement non-linéaire où la performance de tous les scénarios est détériorée. Ce qui est important à noter est le fait que le schéma Pol-Mux OFDM codé ne provoque pas des pénalités supplémentaires en présence des effets non-linéaires de la fibre. Cette validation expérimentale confirme une précédente étude numérique du comportement des codes Polarisation-Temps dans le régime de propagation non-linéaire que nous avons réalisés au préalable.

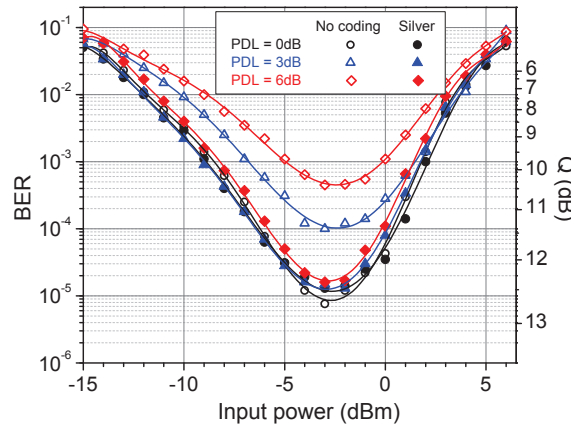


FIGURE E.8: Évolution du TEB en fonction de la puissance injectée après $5 \times 200\text{km}$ pour le cas non-codé et pour le Silver code, à trois niveaux de PDL différents : 0, 3 et 6dB.

E.2.3.3 Compensation des pénalités dans le cas d'une PDL distribuée

Après avoir remis l'émulateur de PDL de 2dB dans la boucle, nous effectuons 2000 mesures de facteur Q pour les différents schémas de codage pour une puissance correspondant au point de fonctionnement optimal. L'OSNR mesuré au niveau du récepteur, et ramené à une bande de 0.1nm, est de 12dB après ajout supplémentaire de bruit ASE afin d'évaluer des distributions statistiques complètes des facteurs Q. Les distributions obtenues sont présentées sur la figure E.9. Nous montrons également, dans la figure insérée, la distribution de probabilité de la PDL ainsi que la distribution théorique à laquelle on s'attend (Maxwellienne de moyenne environ $2 \times \sqrt{5} \approx 4.5\text{dB}$ [53]). La PDL est calculée à partir des matrices de passage 2×2 du canal optique estimées en utilisant la séquence d'apprentissage de la trame OFDM.

La distribution du facteur Q pour les sous-porteuses portant des symboles 4-QAM est asymétrique et s'étale davantage vers les mauvais facteurs Q. Elle a une moyenne de 11.2dB et un écart type (std pour standard deviation) de 0.55dB. D'autre part, la distribution est plus symétrique et le facteur Q moyen est plus élevé et égal à 11.8dB pour les sous-porteuses codées en

Silver ou en Golden (ce qui correspond à la même valeur observée pour le cas sans PDL pour un OSNR égal à 12dB). Les distributions de facteur Q sont également plus étroites lorsque les codes Polarisation-Temps sont utilisés. Le Silver code donne une distribution légèrement plus étroite que celle obtenue avec le Golden code, les écarts-types étant respectivement de 0.35dB et 0.37dB.

Pour le code d'Alamouti, nous observons un facteur Q moyen de 8.3dB en raison de l'utilisation des symboles 16-QAM pour garantir le même rendement binaire que celui des autres scénarios. Cependant, il offre la distribution de facteur Q la plus étroite de toutes avec un std de 0.24dB. Cela est dû à la structure orthogonale puissante de la matrice des mots de code d'Alamouti qui rend sa performance indépendante du taux accumulé de PDL, comme nous avons remarqué et justifié précédemment en simulations et par l'étude théorique. La petite variance observée de la distribution du facteur Q peut être attribuée à la sensibilité accrue de la modulation 16-QAM au bruit de phase ou aux effets non-linéaires.

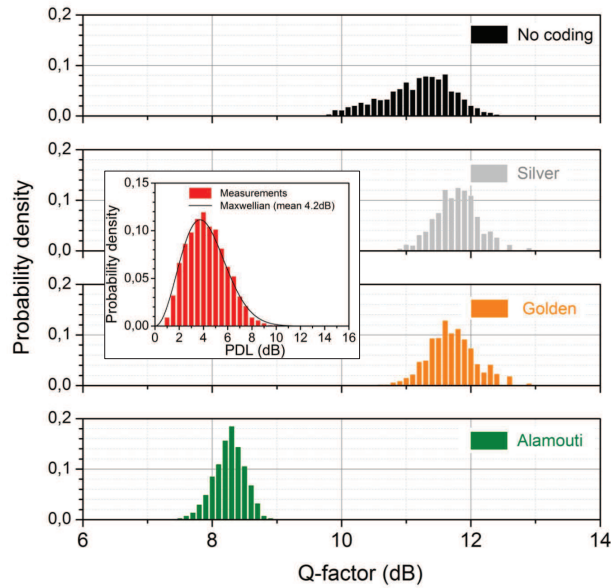


FIGURE E.9: Distributions des facteurs Q après une transmission de $5 \times 200\text{km}$ à $P_{in} = -3\text{dBm}$ ($\text{OSNR}_{0.1\text{nm}} = 12\text{dB}$). Insertion : Distributions théorique et expérimentale de la PDL.

E.2.4 Conclusion

Nous avons présenté une étude numérique et une analyse théorique de la réduction des pénalités causées par la PDL dans les systèmes de transmission optique à multiplexage en polarisation, en utilisant des schémas de codage MIMO qui n'introduisent aucune redondance. Le Silver code peut efficacement réduire les effets de PDL dans les liens optiques de très longue distance où les valeurs de PDL dépassent rarement 6dB [52]. Cependant, l'étude théorique de la compensation de la PDL a montré que ce code n'est pas optimal. En effet, en calculant la borne supérieure de la

probabilité d'erreur pour un canal optique avec PDL, nous avons obtenu le critère de construction de codes qui garantissent une probabilité d'erreur indépendante de la PDL, restaurant ainsi la performance d'un canal à bruit additif blanc Gaussien. Ces critères ouvrent la voie à la construction de codes optimaux spécifiques à la compensation de la PDL dans les canaux optiques. De plus, une expérience de transmission optique dans un lien où la PDL est distribuée, a été aussi réalisée, validant les résultats théoriques obtenus avec un modèle de canal global et simplifié. En outre, nous avons montré que le codage Polarisation-Temps n'introduit pas de pénalités supplémentaires en présence des effets non-linéaires et que ce codage améliore à la fois la moyenne et la variance des distributions du facteur Q quand la PDL est distribuée.

Dans la section suivante, nous nous intéresserons à des systèmes MIMO optiques de plus grandes dimensions : les systèmes utilisant le multiplexage spatial dans les fibres multi-modes, pour explorer la capacité du codage spatio-temporel à réduire les pénalités induites par des effets non-unitaires connus sous le nom de MDL ou Mode Dependent Loss (perte dépendant du mode spatial de propagation).

E.3 Codage Espace-Temps pour les systèmes de transmission optique SDM-OFDM

Pour faire face à la demande toujours croissante de débits d'information plus élevés, un degré de liberté supplémentaire dans la fibre : l'espace est largement étudié aujourd'hui pour réaliser le prochain saut significatif en termes de capacité [1, 13, 82]. Le multiplexage spatial (SDM) a été démontré expérimentalement à travers différentes approches : fibres multi-cœur (MCF) où plusieurs cœurs couplés ou découplés coexistent dans la même gaine, fibres multi-mode (MMF) où le cœur de la fibre est agrandie pour permettre la propagation de plusieurs modes de propagation polarisés linéairement (modes LPs), ou même une combinaison des deux (MCF où chaque cœur est multi-mode) [1, 10, 66, 176]. Le principal avantage de ces techniques SDM est la fourniture d'un ensemble orthogonal de voies qui peuvent transporter des flux de données indépendants dans la même fibre tout en répondant à des contraintes d'intégration des composants du système, à savoir les composants qui fonctionnent sur tous les modes de transport tels que des coupleurs de mode [89], des amplificateurs optiques [68, 92], et des commutateurs sélectifs en longueur d'onde (WSS), visant ainsi à réduire le coût par bit et la consommation d'énergie du système.

Cependant, tout en offrant des capacités de multiplexage intéressantes, le SDM ajoute un défi primaire qui consiste à gérer la diaphonie causée par la coexistence de plusieurs canaux dans le même espace qui est celui de la fibre. Pour les MCFs, la diaphonie augmente lorsque la distance inter-cœur est de plus en plus petite tandis que pour les MMFs, les modes LPs se chevauchent

de façon plus significative dans le même noyau, et les données transportées se mélangent aléatoirement lors de la propagation. Les modes LPs observeront aussi des pertes/gains différentiels menant à une disparité de gain/perte modale. Dans cette section, nous nous concentrons sur la gestion des disparités de perte dans les systèmes optiques utilisant les modes LPs pour réaliser du multiplexage spatiale, que nous appellerons systèmes MDM pour Mode Division Multiplexing. Ce choix est justifié, d'une part, par l'émergence de nouveaux procédés de fabrication de composants optiques supportant un nombre limité de modes appelés "few-mode" tels que des fibres few-mode (FMF) [10] et des amplificateurs few-mode (FMA). Les fibres MMFs conventionnelles ayant un grand diamètre de cœur permettent la propagation d'un grand nombre de modes spatiaux avec des retards de propagation inter-modaux importants, les rendant impropres à l'utilisation pour le moment pour les transmissions optiques de longue distance. Les derniers travaux de recherche ont ainsi considéré des fibres FMFs ayant de faibles valeurs de dispersion modale et supportant la propagation de 3, 6 ou même 10 modes spatiaux [82, 83, 176].

Une autre raison principale qui justifie l'intérêt actuel pour les systèmes MDM est l'intégration du DSP et des techniques de détection cohérente pour gérer au niveau du récepteur la diaphonie modale à l'aide de filtres égaliseurs MIMO. Dans le cas où les multiples flux de données restent orthogonaux et conservent leur énergie totale après la propagation dans le canal, ce passage serait équivalent à une transformation unitaire, et une simple opération d'inversion du canal accompagnée d'une élimination de l'interférence entre symboles induite par la dispersion serait suffisante pour séparer les flux et récupérer la performance d'un canal AWGN exempt de diaphonie. Cependant, les systèmes MDM sont impactés par des disparités modales de perte, aussi connu sous le nom de pertes dépendant du mode spatial ou MDL, résultant des imperfections dans le lien tels que des mélanges non-unitaires dans la fibre à cause des pertes par courbure et des épissures [90] ainsi que des composants optiques dont les pertes/gains dépendent des modes incidents [68, 92]. La MDL détériore la performance globale du système [2, 90] parce que les flux de données transportés par les modes se mélangent tout en observant des niveaux d'atténuation différents, et perdent donc leur orthogonalité. Contrairement à la dispersion et à la diaphonie unitaire, la MDL est un effet qui peut limiter la capacité du système MIMO en réduisant le nombre effectif de canaux disponibles pour le multiplexage [64, 74]. Par conséquent, la MDL est certainement un défi qui doit être surmonté afin d'envisager l'implémentation de systèmes MDM comme une solution possible pour l'industrie des télécommunications du futur.

Afin de réduire la MDL dans les liens optiques de très longue distance où des amplis optiques FMA seront régulièrement insérés pour compenser l'atténuation de la fibre, des solutions optiques tels que des fibres avec des modes fortement couplés [2] et des brouilleurs de mode [90, 94] ont été suggérées pour moyenniser les pertes subies par chaque mode, réduisant ainsi la MDL accumulée. Dans [74, 84], les auteurs ont montré le rôle du fort couplage des modes dans la réduction, à

la fois, de la dispersion inter-modale et de la MDL. Bien que ces solutions réduisent la MDL, ils ne sont pas capables de l'éliminer complètement. En outre, certaines exigences strictes sont nécessaires afin d'obtenir la réduction souhaitée de MDL (fortes contraintes sur les paramètres opto-géométriques des fibres, le nombre de brouilleurs de mode dans le lien optique...). D'autre part, dans [2], Lobato *et al.* ont proposé d'utiliser le décodage ML, implémenté à travers une recherche exhaustive, au lieu d'une détection avec forçage à zéro (ZF) pour améliorer la performance des systèmes MDM impactés par la MDL. Cette amélioration est attendue à coup sûr car, en présence de la MDL, l'opération d'inversion du canal réalisée par l'égalisation ZF amplifiera le bruit reçu et mènera ainsi à une détection sous-optimale. Dans leurs travaux ultérieurs [94], les mêmes auteurs ont proposé deux alternatives à la recherche exhaustive, dont la complexité croît rapidement avec la taille du système MIMO, pour implémenter la détection ML : une version sous-optimale quasi-ML à complexité réduite ainsi qu'un décodeur en réseaux de point, en particulier le célèbre décodeur par sphères [155]. Cependant, nous allons montrer que davantage de gain en performance peut être obtenu en utilisant du codage MIMO plus élaboré qu'un simple multiplexage des données pour atténuer les pénalités de la MDL.

Dans ce qui suit, nous proposons des solutions de codage spatio-temporel (ST) pour combattre la MDL et assouplir les exigences de disparités modales de gain dans les composants optiques conçus pour les systèmes MDM. Nous associons de nouveau le format multi-porteuses OFDM aux codes ST pour faire face aux effets de dispersion et pour pouvoir représenter la transmission, dans le domaine fréquentiel, par un modèle de canal linéaire à un étage unique par sous-porteuse, rendant ainsi le décodage des codes ST moins complexes. L'association des codes ST et de l'OFDM dans les systèmes MIMO multi-antenne doit son succès principalement pour cette raison. Nous montrerons que des schémas ST sans redondance peuvent être appliqués aux systèmes MDM pour améliorer les performances tout en conservant les gains de multiplexage. Les performances des systèmes MDM, de 3 et de 6 modes, comportant des amplis optiques FMAs considérées parmi les composants qui introduisent le plus de MDL [92], seront étudiées sans et avec des codes ST :

- D'abord, nous définissons un modèle de canal MDM et évaluons les pénalités induites par la MDL pour différents taux de couplage modal, et dans un contexte mono-polarisation. Un schéma mono-polarisation est envisagée afin de se concentrer uniquement sur les disparités de perte entre les modes spatiaux.
- En second lieu, nous appliquerons les codes ST à l'émetteur et y associons une détection optimale ML au niveau du récepteur pour combattre la MDL. Nous montrerons, pour la première fois, une compensation complète des pénalités engendrées par des niveaux aussi élevés que 10dB de MDL dans un système MDM de 6 modes.

- Ensuite, nous analysons la complexité de décodage des solutions ST proposées afin d'évaluer les possibilités de mise à l'échelle aux systèmes MDM de plus grande dimension. En conséquence, nous suggérons deux schémas ST de faible complexité, flexibles et extensibles. Le premier schéma consiste à utiliser une détection sous-optimale : le forçage à zéro avec retour de décision (ZF-DFE pour Zero Forcing with Decision Feedback Equalizer) tout en gardant la pénalité induite par la MDL faible. Le second schéma consiste en une nouvelle approche de codage en sous-blocs qui offre un potentiel important pour étendre les solutions ST aux systèmes MDM de plus grandes dimensions.

E.3.1 Systèmes de transmission optique multiplexés en modes spatiaux

En se basant sur des modèles de canaux proposés dans des travaux précédents [2, 78, 90], nous définissons et simulons un modèle de canal MDM impacté par la MDL dans divers scénarios de couplage inter-modal.

E.3.1.1 Modélisation du couplage dans la fibre et les sources de MDL

En se propageant, les modes se chevauchent spatialement dans la fibre en raison des imperfections dans le dopage et des asymétries géométriques causées par un stress mécanique et/ou thermique et des micro-courbures de la fibre. Par conséquent, les symboles portés par chaque mode se mélangent de façon aléatoire tout au long du lien optique. Ce couplage peut être unitaire et sans pertes [76, 77] ou non-unitaire et avec pertes [78] selon les paramètres opto-géométriques et les conditions d'installation des fibres en question ainsi que des autres composants optiques multi-mode d'un lien MDM. Deux stratégies de couplage inter-modal ont été récemment étudiées : tandis que certains travaux ont envisagé de réduire le couplage inter-modal dans les FMFs [79, 80], en maintenant les niveaux de diaphonie faibles pour éviter l'utilisation d'un traitement MIMO complexe au niveau du récepteur, d'autres ont estimé que le couplage serait de toute façon inévitable (surtout dans les composants optiques discrets : Mux/DeMux, amplis...) et qu'un traitement MIMO complet serait donc nécessaire [83]. En outre, Kahn *et al.* [74, 84] ont suggéré d'exploiter positivement le couplage fort vu que dans ce cas, la diffusion maximale des délais de propagation inter-modaux ainsi que le MDL accumulée suivront un processus de marche aléatoire, et évolueront donc en fonction de la racine carrée de la longueur de la fibre ou du nombre de sources de MDL dans le lien, plutôt que d'évoluer linéairement dans le cas de couplage inter-modale faible.

Dans notre modèle de canal, nous choisissons volontairement de garder le couplage généré dans la fibre unitaire, négligeant ainsi toute MDL provenant des épissures imparfaites ou des micro-courbures. Ceci est fait pour se concentrer uniquement sur la MDL accumulée générée par les amplificateurs optiques FMAs, difficile à annuler par des techniques optiques simples en raison des limites technologiques (profils de dopage sur mesure dans les amplificateurs optiques,

la distribution spatiale de la puissance de pompage ...) [92]. Sachant que les amplificateurs optiques sont régulièrement insérés dans une liaison optique de longue distance, les disparités de gain conduiront à l'apparition de valeurs élevées de MDL à la fin du lien.

Pour modéliser le couplage de mode distribué, chaque tronçon de fibre sera représenté par une concaténation de K sections indépendantes :

$$\mathbf{F}_{span,M \times M} = \prod_{k=1}^K (\mathbf{T}_k \mathbf{R}_k) \quad (\text{E.27})$$

Chaque section est modélisée par le produit d'une matrice diagonale $\mathbf{T}_k$ avec des éléments de phase diagonaux tirés uniformément dans $[0 : 2\pi]$ modélisant les décalages de phase entre les modes de propagation, et une matrice de rotation réelle $\mathbf{R}_k$ représentant la diaphonie modale distribuée. Afin de générer ces matrices de rotation pour un certain nombre de modes, les angles de couplage de mode de $\mathbf{R}_k$ seront calculés en utilisant les niveaux de diaphonie générés au niveau d'un désalignement entre les cœurs de deux tronçons de fibres consécutifs. Ces niveaux de diaphonie sont calculées par une intégrale de recouvrement des distributions des champs électriques des modes de propagation sur la section transversale de la fibre à l'endroit des cœurs décalés [90] :

$$c_{M1-M2} = \iint_A E_{M1}(x,y) E_{M2}^*(x+\Delta x, y+\Delta y) dA \quad (\text{E.28})$$

où $M1$ et $M2 \neq M1$ sont deux modes guidés différents, A est la section transversale de la fibre, $E(x,y)$ est la valeur du champ électrique du mode au point (x,y) , Δx et Δy sont les valeurs de désalignement. Il est important de mentionner que nous utilisons les intégrales de recouvrement à l'endroit d'un désalignement fictif comme outil pour émuler le couplage inter-modal. Dans le cas où un véritable désalignement était présent, la matrice de diaphonie résultante ne serait pas unitaire [90]. Cette technique d'émulation a été motivée par une observation dans [180] où les auteurs ont noté que les niveaux de diaphonie calculés à partir des intégrales de recouvrement sont les mêmes que ceux calculés à partir de la théorie des modes couplés (CMT) qui modélise mieux les origines physiques du couplage. Dans la CMT, les coefficients de couplage dépendent de la variation aléatoire de la position du centre du cœur ainsi que de la différence entre les constantes de propagation des modes [76]. Toutefois, la simulation de la CMT pour un grand nombre de modes est gourmand en ressources et moins pratique que la modélisation proposée. $M(M-1)/2$ valeurs de diaphonie sont nécessaires pour constituer la matrice de rotation réelle, ce qui correspond au nombre de ses angles de rotation. Par exemple, pour un système de 3 modes utilisant les mode spatiaux LP_{01} et les deux modes dégénérés du LP_{11} , chaque valeur de diaphonie $c_{LP_{01}-LP_{11a}}$,

$c_{LP_{01}-LP_{11b}}, c_{LP_{11a}-LP_{11b}}$ est équivalente à un angle d'Euler de la matrice de rotation :

$$\mathbf{R}_{3 \times 3}(\phi, \theta, \psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \cdot \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \cdot \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{E.29})$$

où $\phi = \arcsin c_{LP_{11a}-LP_{11b}}$, $\theta = \arcsin c_{LP_{01}-LP_{11b}}$ et $\psi = \arcsin c_{LP_{01}-LP_{11a}}$. Le calcul de matrices de diaphonie unitaire pour 6 modes est une généralisation de cette procédure mettant en jeu 15 valeurs de diaphonie. Les désalignements fictifs Δx et Δy pour chaque section sont tirés uniformément dans l'intervalle $[-\sigma r_c : \sigma r_c]$ où σ est un pourcentage du rayon du cœur r_c déterminant le taux de couplage pour un nombre de section K donné.

E.3.1.2 Statistiques de la MDL dans un système MDM de longue distance

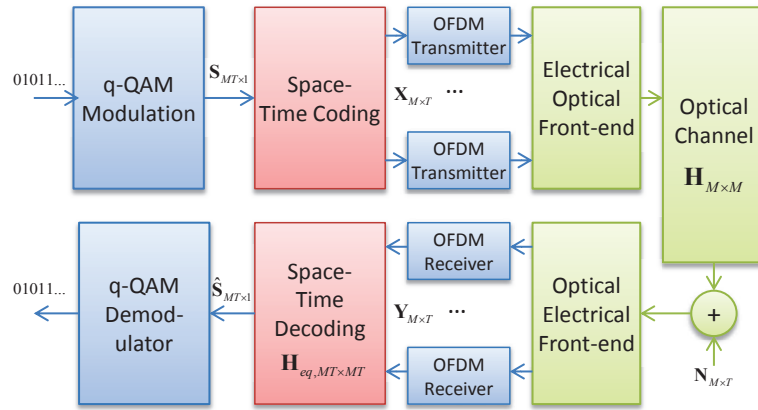


FIGURE E.10: Système de transmission optique OFDM multiplexé en modes, et codé en espace et en temps.

Après avoir présenté les principaux effets linéaires du canal optique, nous définissons un système de transmission OFDM avec 3 (resp. 6) modes spatiaux, utilisés en mode mono-polarisation. Les modes sont le mode fondamental LP_{01} et les deux dégénérescences LP_{11a} et LP_{11b} du mode LP_{11} (resp. ces trois modes ainsi que le LP_{02} ainsi que les dégénérescences LP_{21a} et LP_{21b} du mode LP_{21}). Même si elle est irréaliste, l'hypothèse d'utilisation d'une polarisation par mode spatial est choisi afin de se concentrer uniquement sur les pertes différentielles inter-modales comme dans le cas de plusieurs études précédentes sur la MDL [74, 78, 90]. Nous avons également négligé les non-linéarités de la fibre pour se concentrer sur les effets linéaires du système pour évaluer sa performance. Un signal OFDM avec un préfixe cyclique de taille appropriée est utilisé pour

moduler chacun des modes qui se propagent à travers une liaison optique de longue distance comportant des amplificateurs FMAs avec des disparités de gains modaux. Le schéma de transmission est illustré sur la figure E.10. En l'absence de bruit de phase du laser et de décalage de fréquence, le canal MIMO pour chaque sous-porteuse OFDM peut être exprimé par :

$$\begin{aligned} \mathbf{Y}_{M \times T} &= \mathbf{H}_{M \times M} \mathbf{X}_{M \times T} + \mathbf{N}_{M \times T} \\ &= \sqrt{\alpha} \prod_{l=1}^L (\mathbf{P}_l \mathbf{G}_l \mathbf{F}_l) \mathbf{X}_{M \times T} + \mathbf{N}_{M \times T} \end{aligned} \quad (\text{E.30})$$

où $\mathbf{X}_{M \times T}$ (resp. $\mathbf{Y}_{M \times T}$) sont les symboles complexes émis (resp. reçus) sur les $M = \{3, 6\}$ modes et au cours de T créneaux temporels. $\mathbf{H}_{M \times M}$ est la matrice de canal linéaire constitué de L tronçons de fibre indépendants $\mathbf{F}_l$, chacun étant de la forme (E.27) et suivi par un FMA modélisé comme une matrice diagonale $\mathbf{G}_l$ et par un brouilleur de modes $\mathbf{P}_l$. Les gains dans $\mathbf{G}_l$ sont attribuées comme suit : le mode LP_{01} a un gain unitaire et des décalages des gains $\Delta G_{01-\mu\nu}$ sont définis pour chaque mode $LP_{\mu\nu}$. $\mathbf{P}_l$ sont des matrices de permutation aléatoires obtenus en permutant aléatoirement les lignes d'une matrice d'identité $\mathbf{I}_M$, représentant ainsi des brouilleurs de modes parfaits comme dans [90]. α est un facteur de normalisation introduit pour compenser la perte du lien commune à tous les modes. Dans un système réel, la MDL serait dépendante de la fréquence avec une bande de cohérence dictée par la quantité de dispersion modale dans la liaison [74]. En première approche, la dispersion modale n'a pas été prise en considération, pour limiter l'étude en cours à une seule sous-porteuse OFDM ou à un canal plutôt plat en fréquence.

Avant d'introduire les codes ST, nous examinons les statistiques de la MDL dans le modèle canal défini. Pour évaluer la MDL sous différents scénarios de couplage, nous considérons un système MDM avec $L = 8$ tronçons de fibres FMF ayant un profil d'indice parabolique avec un rayon de cœur de $8.7\mu\text{m}$ et une ouverture numérique de 0.205 à $\lambda = 1550\text{nm}$, permettant ainsi la propagation de 6 modes spatiaux. Chaque tronçon est constitué de $K = 200$ sections. Les amplificateurs présentent les disparités de gain suivantes $\Delta G_{01-11} = -1.3\text{dB}$, $\Delta G_{01-21} = -2\text{dB}$ et $\Delta G_{01-02} = -0.2\text{dB}$ correspondant à une technologie prometteuse de FMA présentée dans [92], et ils sont suivis par un brouilleur de mode parfait. Pour $K = 200$ sections, trois taux de couplage sont examinés en tirant aléatoirement des désalignements d'une distribution uniforme avec σ accordé à 0.6%, 3% et 5% du rayon du cœur de la fibre r_c pour émuler respectivement du faible, du moyen et du fort couplage inter-modal. Pour chaque scénario de couplage, 10^6 états du canal MDM 6×6 ont été tirés aléatoirement et les niveaux de MDL ont été calculés pour chacun, la MDL étant défini comme le rapport en dB entre les carrés de la plus grande et de la plus petite des valeurs singulières du canal $\mathbf{H}$. Les distributions de la MDL sont présentées sur la figure E.11 pour les différents niveaux de couplage, avec et sans l'option de brouillage de mode.

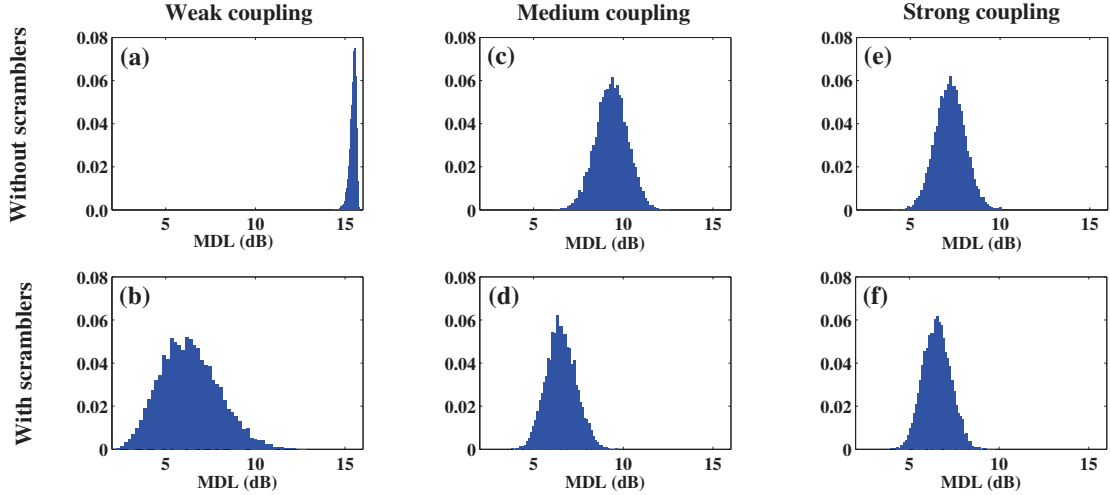


FIGURE E.11: Distributions de probabilité de la MDL simulées pour différents scénarios de couplage et de brouillage de modes dans un système MDM 6×6 (8 tronçons, 2dB de MDL par tronçons).

Nous remarquons clairement que le couplage et le brouillage réduisent de manière significative la MDL moyenne ainsi que sa variance. Avec des tronçons à faible couplage, les 2dB de MDL de chaque amplificateur s'additionnent aboutissant à une MDL cumulée égale à $8 \times 2 = 16\text{dB}$. A un taux de couplage moyen, les modes ne sont que partiellement corrélés et la MDL moyenne diminue jusqu'à 10dB. Le fort couplage et le brouillage des modes réduisent la MDL moyenne à 6dB, ce qui est très proche de $\sqrt{8} \times 2 = 5.7\text{dB}$, la valeur de MDL accumulée prévue lorsqu'un mélange ou un couplage aléatoire complet des modes se produit entre des sources de MDL identiques [74]. Par conséquent, un fort couplage dans les fibres FMF est souhaitable afin de réduire la MDL accumulée. Cependant, dans tous les scénarios, nous constatons que la MDL n'est pas complètement éliminée et proposons donc d'ajouter des solutions de codage MIMO pour limiter les pénalités induites par la MDL.

E.3.2 Codage Espace-Temps et décodage en sphères

Comme nous l'avons vu précédemment avec les systèmes MIMO 2×2 , nous pouvons faire un meilleur usage des dimensions du système qu'un simple multiplexage de données que nous appliquerons aux systèmes MDM, en insérant par exemple le même symbole de données dans des combinaisons linéaires envoyées sur différents modes et à différents créneaux temporels. Par la suite, au niveau du récepteur, les copies d'une même donnée peuvent être exploitées pour améliorer la performance car le même symbole aurait connu divers états du canal, et par conséquent les disparités de perte entre les flux de données seraient davantage réduites.

E.3.2.1 Les codes ST algébriques en couches (TAST)

Plusieurs familles de codes ST ont été conçues pour différents schémas MIMO sans fils. Nous mettrons l'accent sur une catégorie spécifique de codes : les codes espace-temps en blocs (STBC ou Space-Time Block codes) dans laquelle un mot de code est représenté par une matrice $\mathbf{X}_{M \times T}$, où T est le nombre de créneaux temporels sur lesquels le code est défini. Cette matrice est obtenue en multipliant un vecteur de symbole $\mathbf{S}_{MT \times 1}$ par une matrice génératrice du code $\mathbf{M}_G$ et en réarrangeant le vecteur obtenu en une matrice $M \times T$. Cette opération a lieu du côté émetteur dans le bloc "Space-Time coding" dans la figure E.10.

Dans cette catégorie, nous choisirons des codes pour les systèmes MDM 3×3 et 6×6 , qui remplissent les conditions suivantes : en premier, les matrices de mots de code sont carrées ($M = T$) afin de placer chaque symbole de données sur un mode différent à chaque créneau temporel différent, et pour minimiser en même temps la latence au niveau du récepteur. En second, des codes à rendement plein seront utilisés, ce qui se traduit par la transmission de M^2 symboles d'information q-QAM différents dans chaque mot de code pour ne pas réduire le gain de multiplexage du système MDM. En troisième lieu, une énergie moyenne uniforme sera transmise par mode. Une famille de codes ST en blocs connue sous le nom de codes TAST pour Threaded Algebraic Space-Time codes [148], ou les codes ST algébriques en couches, répond à toutes ces exigences et offre des codes pour tout système MIMO $M \times T$. Les matrices de mots de code 3×3 et 6×6 sont données ci-dessous :

- Le mot de code TAST 3×3 est donné en (E.31) où $\phi = \exp(i\pi/36)$, $j = \exp(i2\pi/3)$, $\theta = \exp(i\pi/9)$ et $s_{k=1:9}$ sont des symboles q-QAM. ϕ et θ ont été choisies de telle sorte à maximiser le gain de codage et le gain de diversité sur un canal multi-antenne 3×3 sans-fils à évanouissement de type Rayleigh [148]. 9 symboles q-QAM sont envoyés pendant 3 créneaux temporels (ou time slot (ts)) dans chaque mot de code, fournissant ainsi un débit de 3 symboles/ts. Une autre représentation du code consiste à vectoriser (E.31), en tassant ses colonnes pour obtenir (E.32) où $\mathbf{M}_G$ est la matrice génératrice du code. $\mathbf{M}_G$ étant unitaire ($\mathbf{M}_G \mathbf{M}_G^\dagger = \mathbf{I}_{9 \times 9}$), nous ne modifions alors pas l'énergie moyenne des nouveaux super-symboles transmis après avoir encodé les symboles q-QAM.

$$\mathbf{X}_{T,3 \times 3} = \frac{1}{\sqrt{3}} \begin{bmatrix} s_1 + \theta s_2 + \theta^2 s_3 & \phi^2(s_4 + j\theta s_5 + j^2\theta^2 s_6) & \phi(s_7 + j^2\theta s_8 + j\theta^2 s_9) \\ \phi(s_7 + \theta s_8 + \theta^2 s_9) & s_1 + j\theta s_2 + j^2\theta^2 s_3 & \phi^2(s_4 + j^2\theta s_5 + j\theta^2 s_6) \\ \phi^2(s_4 + \theta s_5 + \theta^2 s_6) & \phi(s_7 + j\theta s_8 + j^2\theta^2 s_9) & s_1 + j^2\theta s_2 + j\theta^2 s_3 \end{bmatrix} \quad (\text{E.31})$$

$$\begin{aligned} \text{vec}_{\mathbb{C}}(\mathbf{X}_{\mathcal{T},3 \times 3}) &= \mathbf{M}_{G,9 \times 9} \mathbf{S}_{9 \times 1} \\ &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & \theta & \theta^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \phi\theta & \phi\theta^2 \\ 0 & 0 & 0 & \phi^2 & \phi^2\theta & \phi^2\theta^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi^2 & \phi^2j\theta & \phi^2j^2\theta^2 & 0 & 0 & 0 \\ 1 & j\theta & j^2\theta^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \phi j\theta & \phi j^2\theta^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & \phi j^2\theta & \phi j\theta^2 \\ 0 & 0 & 0 & \phi^2 & \phi^2j^2\theta & \phi^2j\theta^2 & 0 & 0 & 0 \\ 1 & j^2\theta & j\theta^2 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \end{bmatrix} \end{aligned} \quad (\text{E.32})$$

- La matrice de mot de code TAST 6×6 est donnée en (E.33) où $\phi = \exp(i\pi/12)$, $\mathbf{s}_{1 \dots 6}$ un vecteur de 6 symboles q-QAM, $f_n(\mathbf{x}) = \sum_{k=1:6} x_k (j^{n-1}\theta)^{k-1}$ avec $j = \exp(i2\pi/6)$ et $\theta = \exp(i\pi/18)$. ϕ et θ ont été fixées afin de maximiser les gains de codage et de diversité sur un canal sans fils multi-antenne 6×6 à évanouissement de type Rayleigh [148]. 36 symboles q-QAM sont transmis durant 6 créneaux temporels dans chaque mot de code, garantissant ainsi un rendement plein de 6 symboles/ts. La matrice génératrice correspondant qui est aussi unitaire peut être aussi construite par vectorisation de (E.33).

$$\mathbf{X}_{\mathcal{T},6 \times 6} = \frac{1}{\sqrt{6}} \begin{bmatrix} f_1(\mathbf{s}_1) & \phi^{\frac{5}{6}}f_2(\mathbf{s}_6) & \phi^{\frac{4}{6}}f_3(\mathbf{s}_5) & \phi^{\frac{3}{6}}f_4(\mathbf{s}_4) & \phi^{\frac{2}{6}}f_5(\mathbf{s}_3) & \phi^{\frac{1}{6}}f_6(\mathbf{s}_2) \\ \phi^{\frac{1}{6}}f_1(\mathbf{s}_2) & f_2(\mathbf{s}_1) & \phi^{\frac{5}{6}}f_3(\mathbf{s}_6) & \phi^{\frac{4}{6}}f_4(\mathbf{s}_5) & \phi^{\frac{3}{6}}f_5(\mathbf{s}_4) & \phi^{\frac{2}{6}}f_6(\mathbf{s}_3) \\ \phi^{\frac{2}{6}}f_1(\mathbf{s}_3) & \phi^{\frac{1}{6}}f_2(\mathbf{s}_2) & f_3(\mathbf{s}_1) & \phi^{\frac{5}{6}}f_4(\mathbf{s}_6) & \phi^{\frac{4}{6}}f_5(\mathbf{s}_5) & \phi^{\frac{3}{6}}f_6(\mathbf{s}_4) \\ \phi^{\frac{3}{6}}f_1(\mathbf{s}_4) & \phi^{\frac{2}{6}}f_2(\mathbf{s}_3) & \phi^{\frac{1}{6}}f_3(\mathbf{s}_2) & f_4(\mathbf{s}_1) & \phi^{\frac{5}{6}}f_5(\mathbf{s}_6) & \phi^{\frac{4}{6}}f_6(\mathbf{s}_5) \\ \phi^{\frac{4}{6}}f_1(\mathbf{s}_5) & \phi^{\frac{3}{6}}f_2(\mathbf{s}_4) & \phi^{\frac{2}{6}}f_3(\mathbf{s}_3) & \phi^{\frac{1}{6}}f_4(\mathbf{s}_2) & f_5(\mathbf{s}_1) & \phi^{\frac{5}{6}}f_6(\mathbf{s}_6) \\ \phi^{\frac{5}{6}}f_1(\mathbf{s}_6) & \phi^{\frac{4}{6}}f_2(\mathbf{s}_5) & \phi^{\frac{3}{6}}f_3(\mathbf{s}_4) & \phi^{\frac{2}{6}}f_4(\mathbf{s}_3) & \phi^{\frac{1}{6}}f_5(\mathbf{s}_2) & f_6(\mathbf{s}_1) \end{bmatrix} \quad (\text{E.33})$$

E.3.2.2 Décodeurs ML pour les systèmes MIMO de grande dimension

Du côté du récepteur, les symboles de données sont estimés en utilisant un décodeur de maximum de vraisemblance (ML). Supposons que la matrice canal $\mathbf{H}$ est connue (ou parfaitement estimée au niveau du récepteur) et constante pendant T temps symbole, et que tous les mots de code émis $\mathbf{X}$ sont équiprobables, un système de détection optimale du canal modélisé en (E.30) doit satisfaire au critère ML qui consiste à estimer le mot de code $\mathbf{X}$ par $\hat{\mathbf{X}}_{ML}$ en minimisant la distance Euclidienne suivante :

$$\hat{\mathbf{X}}_{ML} = \underset{\mathbf{X}_{M \times T} \in \mathcal{C}}{\text{argmin}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 \quad (\text{E.34})$$

où $\mathcal{C}$ est l'ensemble de tous les mots de code possibles. Le critère ML peut être réécrit pour faire apparaître explicitement les symboles q-QAM originaux lorsqu'un schéma de codage ST est utilisé. À cette fin, nous utilisons une forme vectorisée de (E.30) contenant la matrice génératrice

du code ST et définissons un canal équivalent $\mathbf{H}_{eq}$ [154] :

$$\begin{aligned} \mathbf{Y}'_{MT \times 1} = \text{vec}_{\mathbb{C}}(\mathbf{Y}) &= \begin{bmatrix} \mathbf{H} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H} \end{bmatrix} \text{vec}_{\mathbb{C}}(\mathbf{X}) + \text{vec}_{\mathbb{C}}(\mathbf{N}) \\ &= \mathbf{H}'_{MT \times MT} \mathbf{M}_G \mathbf{S}_{MT \times 1} + \mathbf{N}'_{MT \times 1} \\ &= \mathbf{H}_{eq} \mathbf{S} + \mathbf{N}' \end{aligned} \quad (\text{E.35})$$

où $\mathbf{M}_G$ est la matrice génératrice du schéma de codage. Dans le cas de multiplexage spatial simple, sans codage, $T = 1$ et $\mathbf{M}_G$ est remplacée par la matrice identité. $\mathbf{S}$ est le vecteur des symboles q-QAM transmis. Étant donné que $\mathbf{H}$ est une matrice de rang plein et $\mathbf{M}_G$ est unitaire, la règle de décodage ML peut être réinterprétée comme :

$$\hat{\mathbf{S}}_{ML} = \underset{\mathbf{S}_{MT \times 1} \in \mathcal{C}'}{\text{argmin}} \|\mathbf{Y}' - \mathbf{H}_{eq} \mathbf{S}\|^2 \quad (\text{E.36})$$

où $\mathcal{C}'$ est l'ensemble de tous les vecteurs de symboles q-QAM possibles. Le critère ML peut être mis en œuvre par une recherche exhaustive dans $\mathcal{C}'$. La norme dans (E.36) doit être alors calculée pour toutes les combinaisons possibles des symboles émis. Dans le cas d'un code ST carré, à rendement plein, construit avec des symboles q-QAM, cela conduit à q^{M^2} calculs de norme. Ainsi, la complexité de la recherche exhaustive augmente avec la taille de constellation et de façon exponentielle avec le nombre de modes. Par conséquent, des décodeurs à complexité réduite sont indispensable pour implémenter le critère ML.

Des décodeurs en réseau de points à complexité réduite comme le décodeur par sphères ont été suggérés dans [156] pour les systèmes MIMO non-codés ou codés. Après l'application d'une transformation complexe-réelle du canal donné en (E.35), le dictionnaire des symboles $\mathcal{C}'$ est vue comme un réseau de points où chaque vecteur de symbole émis $\mathbf{S}$ est un point du réseau. En raison de la propagation à travers un canal bruité, le vecteur reçu $\mathbf{Y}'$ n'est plus un point du réseau, et le décodeur par sphères recherche le point du réseau le plus proche de $\mathbf{Y}'$ dans une région sphérique centrée sur $\mathbf{Y}'$. Le choix du rayon initial de la sphère est critique pour la réduction de la complexité de l'algorithme de recherche. Si le rayon est judicieusement choisi en fonction du SNR et des valeurs singulières du canal, la complexité de l'algorithme de recherche devient indépendante de la taille de constellation q et est approximée par $\mathcal{O}((MT)^6)$ [156]. Cela donne $\mathcal{O}(M^6)$ opérations dans le cas du multiplexage spatial non codé, et $\mathcal{O}(M^{12})$ pour un code ST en blocs, carré et de rendement plein.

E.3.3 Performance des systèmes MDM codés en espace et en temps

Les avantages de l'utilisation du codage ST pour compenser la MDL seront illustrées dans un système MDM de 3 modes (resp. de 6 modes) où des fibres à gradient d'indice avec un profil d'indice parabolique sont installées, avec un rayon de cœur $r_c = 6\mu\text{m}$ (resp. $r_c = 8.7\mu\text{m}$). Les distributions de champ des modes sont approchées par des modes Laguerre-Gauss comme dans [90]. Le lien simulé est celui présenté dans (E.30) et les paramètres de la liaison sont les mêmes que ceux utilisés dans la section E.3.1 ($L = 8$ tronçons, composés chacun de $K = 200$ sections avec trois niveaux de couplage inter-modal et des disparités de gain au niveau des FMAs). Pour le système de 3 modes, un décalage de gain $\Delta G_{01-11} = -1.3\text{dB}$ est considéré à chaque FMA. Nous rappelons qu'une seule polarisation par mode est considérée pour se concentrer uniquement sur la MDL.

A l'émetteur, dans le cas non-codé (NC), un vecteur de symboles 4-QAM $S_{m=1:M}$ d'énergie symbole unitaire $E_S = 1$ est envoyé sur les modes, fournissant ainsi un débit de six bits (resp. 12 bits) par créneau temporel. Pour le cas codé, un code TAST 3×3 (resp. 6×6) est utilisé. La performance de ces systèmes est mesurée en termes de courbes de TEB moyen en fonction du rapport signal-sur-bruit symbole $E_S/2N_0$ par mode, au moyen de simulations Monte-Carlo. Un minimum de 100 erreurs binaires est enregistré pour chaque valeur de SNR. Trois scénarios de couplage avec et sans l'option de brouillage de modes après les FMAs sont présentés sur la figure E.12. Au niveau du récepteur, dans tous les scénarios, les symboles de données sont estimés en utilisant un décodeur par sphères.

La première colonne de la figure E.12 montre les résultats pour un système de MDM de 3 modes. En observant les courbes en trait plein avec des marqueurs en forme carrée correspondant au cas non-codé (NC) sans brouilleurs de mode, nous remarquons que la pénalité SNR induite par la MDL à un $\text{TEB} = 10^{-3}$ diminue de 4.2dB pour un faible niveau de couplage, à 1.2dB pour du couplage moyen et 0.5dB pour du couplage fort. L'ajout de brouilleurs de mode après les FMAs (courbes en trait plein avec marqueurs en triangle) réduit ces pénalités 1.5dB pour le couplage faible et 0.4dB pour le couplage moyen alors qu'il n'a pas d'effet en régime de couplage fort parce que les modes sont déjà totalement mélangés dans la fibre. Les mêmes observations peuvent être faites pour le système MDM de 6 modes (deuxième colonne de la figure E.12). Sans brouilleurs, la pénalité SNR à $\text{TEB} = 10^{-3}$ diminue de plus que 6dB pour du couplage faible, à 1.2dB pour le couplage moyen et à 0.4dB pour couplage fort. L'insertion de brouilleurs de mode réduit ces pénalités à 1.7dB pour le couplage faible et 0.4dB pour le couplage moyen. Encore une fois, le brouillage n'a pas d'effet dans le régime de couplage fort parce que les modes ont subi un mélange complet dans la fibre et la MDL ne peut être réduite davantage par les brouilleurs comme on peut le constater à partir des distributions de MDL obtenues et illustrées sur la figure E.11.

D'autre part, lorsque le code TAST 3×3 est utilisé seul (courbes en traits discontinus avec marqueurs carrés sur la figure E.12), son gain dépasse celui apporté par du brouillage de modes

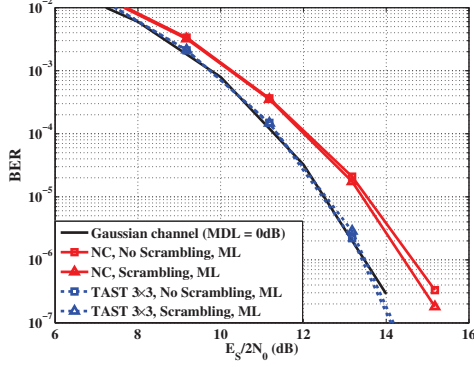
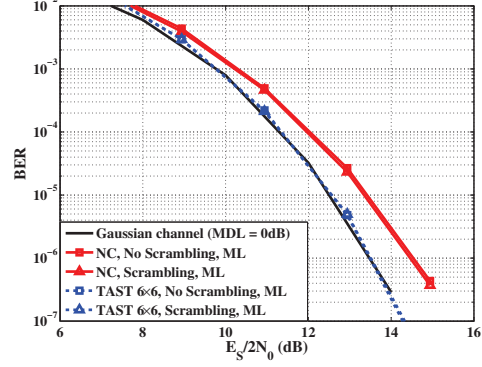
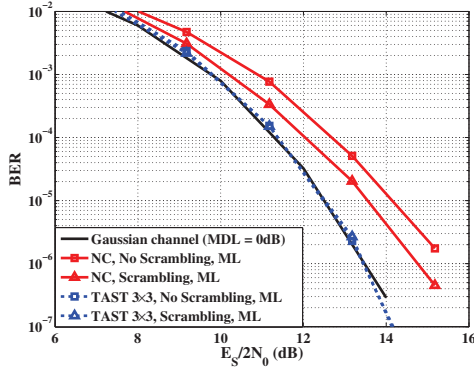
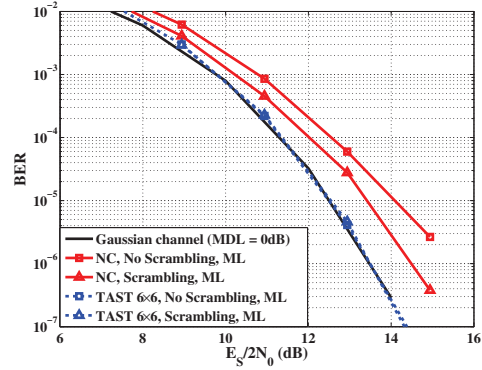
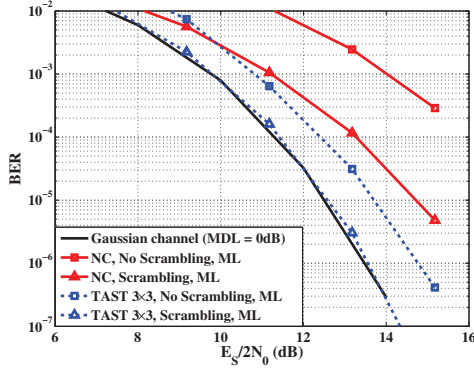
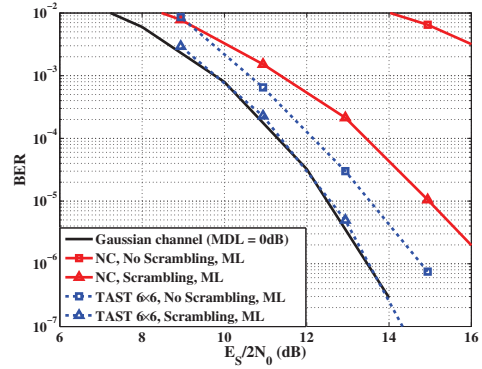

 (a) système 3×3 , fort couplage.

 (b) système 6×6 , fort couplage.

 (c) système 3×3 , couplage moyen.

 (d) système 6×6 , couplage moyen.

 (e) système 3×3 , faible couplage.

 (f) système 6×6 , faible couplage.

FIGURE E.12: TEB moyens en fonction du SNR pour un système MDM 3×3 (a,c,e) et 6×6 (b,d,f) obtenus par des simulations Monte-Carlo (8 tronçons de fibres à gradient d'indice avec un profil d'indice parabolique et un rayon de cœur de $r_c = 6\mu\text{m}$ (resp. $r_c = 8.7\mu\text{m}$) et une ouverture numérique $NA = 0.205$ à $\lambda = 1550\text{nm}$, FMAs avec décalage maximal de gain modal de 2dB).

dans le scénario de faible couplage, et la pénalité SNR est réduite à 1dB pour un $TEB = 10^{-3}$ pour le système de 3 modes, ce qui équivaut à un gain de codage de 3dB. En outre, le code TAST absorbe toute la pénalité due à la MDL dans les scénarios de moyen et de fort couplage. Des réductions similaires des pénalités sont observées avec le code TAST 6×6 pour les différents niveaux de couplage. Enfin, en combinant le code ST avec le brouillage de modes, nous éliminons complètement les pénalités dans tous les régimes de couplage testés pour 3 et 6 modes, et cela à tous les niveaux de TEB que nous avons évalués. Ces résultats prouvent l'efficacité des solutions de codage ST pour combattre la MDL, en moyennant les pertes subies par les symboles de données multiplexés sur les modes ; ce qui en fait une solution DSP intéressante pour les systèmes MDM. A partir des résultats obtenus, nous concluons que les codes ST peuvent être adoptés comme une alternative au brouillage optique des modes ou comme une solution complémentaire pour limiter le nombre de brouilleurs de mode à installer dans la liaison optique. Ce choix entre ses deux solutions se fera en fonction de l'importance des disparités de gain aux FMAs et du taux de couplage dans les fibres FMF utilisées.

Tout comme les codes Polarisation-Temps dans les systèmes PDM où ces codes offraient une plus grande distance minimale que les schémas non-codés, les codes Espace-Temps offrirait le même avantage dans les systèmes MDM impactés par la MDL. Or, la mise à l'échelle des solutions de codage ST pour des systèmes MIMO optiques de plus grandes dimensions devrait garantir une complexité de codage/décodage raisonnable pour que ces solutions puissent être implémentées dans de futurs systèmes MDM.

E.3.4 Complexité et extensibilité des systèmes MDM codés

Les codes TAST à rendement plein précédemment utilisés existent pour des systèmes MIMO de plus grande dimension comme des systèmes MDM de 10 modes ou des systèmes Pol-Mux MDM. Cependant, la complexité de leur décodage augmente rapidement avec la taille du système MIMO : exponentiellement si un décodeur ML avec une recherche exhaustive est utilisé (q^{M^2} pour un code ST carré à base de symboles q-QAM) et en temps polynomial pour un décodeur en sphères (M^{12} pour un code ST carré). Cette complexité accrue peut rendre le coût d'un système MDM codé prohibitif. Par conséquent, nous proposons et étudions deux variantes possibles de solutions de codage ST qui négocient une partie des gains de codage optimaux observés dans la section précédente pour une réduction de la complexité et une possibilité de mise à l'échelle des solutions ST. Les deux stratégies consistent à :

- remplacer le décodeur optimal ML par un égaliseur sous-optimal de plus faible complexité : le forçage à zéro avec retour de décision (ZF-DFE).

- remplacer les codes $M \times M$ par un codage en sous-blocs sur moins de créneaux temporels $T < M$ afin de réduire la complexité du décodeur ML en sphères en raccourcissant la taille du dictionnaire des mots de code.

E.3.4.1 Décodage ZF-DFE des schémas MDM codés

La première solution consiste à utiliser le forçage à zéro avec retour de décision (ZF-DFE) comme égaliseur sous-optimal. Dans les canaux MIMO multi-antenne sans fils à évanouissement de Rayleigh, le ZF-DFE démontre un gain de performance par rapport au décodeur classique ZF qui effectue une simple opération d'inversion du canal, grâce à l'annulation successive de l'interférence lors de la récupération des symboles de données [160]. Les étapes de décodage sont les suivantes :

- Ayant le modèle de canal équivalent en (E.35), effectuer une décomposition QR de $\mathbf{H}_{eq} = \mathbf{Q}\mathbf{R}$ qui réécrit la matrice canal en un produit d'une matrice $\mathbf{Q}$ unitaire et d'une matrice triangulaire supérieure $\mathbf{R}$.
- Calculer le système équivalent :

$$\tilde{\mathbf{Y}} = \mathbf{Q}^\dagger \mathbf{Y}' = \mathbf{R}\mathbf{S} + \mathbf{Q}^\dagger \mathbf{N}' \quad (\text{E.37})$$

où le nouveau bruit $\mathbf{Q}^\dagger \mathbf{N}'$ reste blanc Gaussien vu que $\mathbf{Q}$ est unitaire.

- Estimer chaque symbole du vecteur $\hat{\mathbf{S}}_{ZF-DFE}$ en résolvant le système linéaire $\tilde{\mathbf{Y}} = \mathbf{R}\mathbf{S}$. La matrice $\mathbf{R}$ étant triangulaire supérieure, le système peut être résolu d'une manière itérative en commençant par le dernier symbole et effectuant une décision à seuil sur chaque symbole estimé avant de l'utiliser pour enlever son interférence dans la ligne précédente. Cela se traduit par les opérations suivantes :

$$s_{i,i:N \rightarrow 1} = \frac{y_i - \sum_{j=0}^{i-1} r_{i,N-j} \hat{s}_{N-j}}{r_{i,i}} \quad (\text{E.38})$$

$$\hat{s}_{i,i:N \rightarrow 1} = \lfloor s_i \rfloor$$

où $N = MT$ est la dimension du système et $\lfloor \cdot \rfloor$ est l'opération de décision à seuil.

La complexité de l'algorithme ci-dessus est fixée par la décomposition QR de la matrice canal et la résolution du système linéaire. Nous avons calculé une estimation approximative de cette complexité en flops que nous définissons comme le nombre de multiplications scalaires complexes nécessaires pour décoder un symbole transmis $\mathbf{S}$, et nous l'avons comparé au nombre de flops requis par un décodeur ZF (inversion de matrice en utilisant la méthode du pivot de Gauss standard) ainsi que celui requis par un décodeur ML à recherche exhaustive. Nous avons choisi la recherche

TABLE E.2: Comparaison de la complexité de décodage par symbole de différents décodeurs d'un système MIMO de taille $N \times N$.

	ZF	ZF-DFE	ML exh. search	ML sphere dec.
$N \times N$ MIMO	$\frac{N^3+3N^2-N}{N}$	$\frac{N^3+3.5N^2+0.5N}{N}$	$\frac{q^N(N^2+N)}{N}$	$\approx \frac{N^6}{N}$
6-mode NC ($N = 6$)	53	58	28672	7776
TAST 6×6 ($N = 36$)	1403	1423	1.7×10^{23}	6×10^7

exhaustive pour le décodeur ML parce qu'un nombre exact d'opérations ne peut être obtenue pour un décodeur par sphères. L'expression analytique du nombre de flops nécessaires pour décoder un symbole q-QAM dans un système MIMO $N \times N$ est donnée dans le tableau E.2 pour les différents décodeurs. Un nombre d'opérations approximatif, $\mathcal{O}(N^6)$ [156], est donné pour le décodeur en sphères car le nombre exact dépend du choix du rayon initial et de la valeur de SNR. Une application numérique pour un système MDM de 6 modes non-codé et codé en TAST 6×6 en utilisant des symboles 4-QAM est également indiquée dans le tableau E.2. Il est intéressant de noter que les complexités des décodeurs ZF et ZF-DFE sont dominées par le même terme en N^3 et sont presque égales pour les systèmes MIMO de grandes dimensions.

Nous mesurons la performance du décodeur ZF-DFE avec le système MDM 6×6 précédemment défini. Les courbes de TEB moyen des deux schémas NC et codés en TAST, pour trois taux de couplage inter-modal différents sont données sur la figure E.13. Dans tous les scénarios, les brouilleurs de mode sont utilisés après les FMAs. Tout d'abord, il est évident que le décodeur ZF-DFE offre une performance pire que le décodeur ML pour le régime non-codé (courbes avec marqueurs ronds) en raison de l'amplification du bruit produite lors de l'inversion du canal dans (E.37). D'autre part, lorsque le codage TAST est utilisé avec un décodeur ZF-DFE (courbes en traits discontinus), un gain de performance est observé dans tous les cas. A un TEB = 10^{-3} , le décodage ZF-DFE du schéma codé a la même pénalité de 0.4dB que le schéma non-codé utilisant un décodage ML optimal, pour des taux de couplage moyen et fort, alors qu'il surpasse le système non-codé pour un taux de couplage faible, offrant un gain de codage non-négligeable de 1.2dB. Des gains similaires sont observés pour le système MDM de 3 modes.

Étonnamment, cette performance observée des codes ST est complètement différente de celle obtenue sur un canal sans fils où un décodage ZF-DFE des mots de code n'apporterait absolument aucun gain. Nous avons également testé un décodage ZF des schémas codés qui a montré une performance moins bonne que celle du ZF-DFE. Pour conclure, nous avons montré qu'une solution de codage ST avec un décodage ZF-DFE de faible complexité est également capable d'atténuer les pénalités de la MDL. Dans tous les scénarios de couplage examinés, le décodage sous-optimal

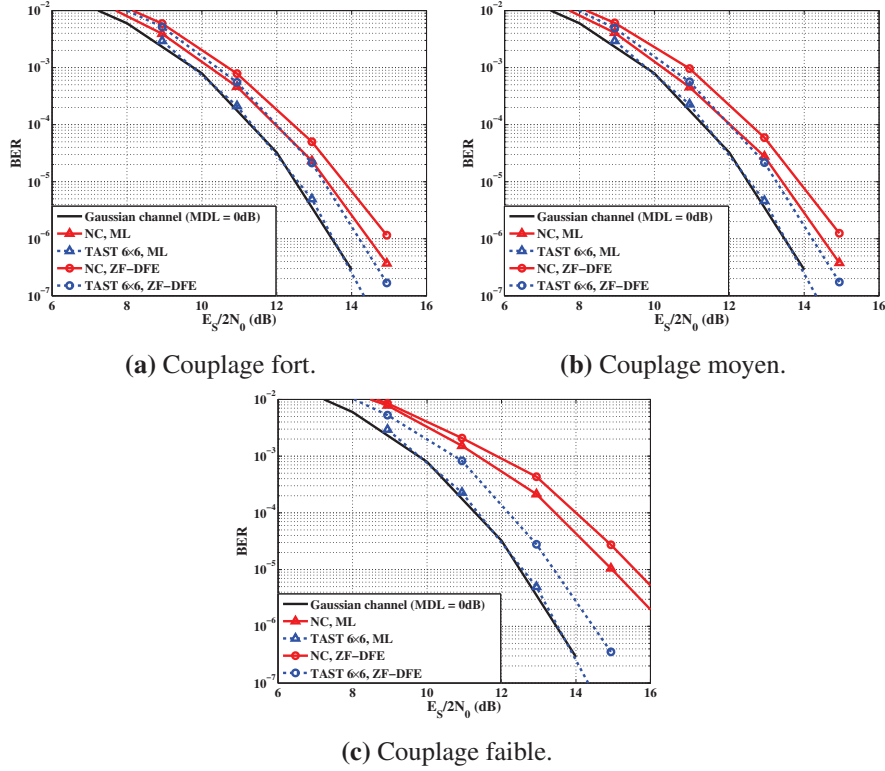


FIGURE E.13: TEB moyens en fonction du SNR du décodage ZF-DFE des schémas codés en TAST 6×6 avec brouillage de modes.

des schémas 6×6 codés, observant des niveaux de MDL aussi élevés que 10 dB, est au pire à 1dB du même schéma décodé d'une façon optimale.

E.3.4.2 Codage spatio-temporel en blocs multiples

La deuxième stratégie consiste en une approche de codage multi-bloc. Encore une fois, nous considérons le précédent système MDM de 6 modes pour étudier deux nouvelles configurations ST à rendement plein et décodé selon le critère ML. Nous suggérons de réduire la dimension du système codé en raccourcissant la longueur temporelle des schémas de codage. Des schémas à rendement plein définis sur $T < 6$ créneaux temporels tels que deux blocs $\mathbf{X}_{3 \times 3}$ codés sur $T = 3$ comme le montre la figure E.14(b) ou trois blocs $\mathbf{X}_{2 \times 2}$ codés sur $T = 2$ comme illustré sur la figure E.14(c) peuvent être envisagés.

Si nous calculons le nombre d'opérations nécessaires pour décoder un symbole dans ces nouveaux schémas ST en appliquant le critère ML à travers une recherche exhaustive, nous constatons que 1.3×10^{12} opérations sont nécessaires quand on considère deux blocs ST sur 3 créneaux temporels et 2.2×10^8 opérations pour trois blocs ST sur 2 créneaux temporels. Une réduction importante de la complexité de décodage est ainsi obtenu par rapport à la solution TAST 6×6 .

En effet, cette dernière nécessite le décodage de mots de code contenant 36 symboles de données alors que les nouveaux schémas contiennent que 18 et 12 symboles de données respectivement.

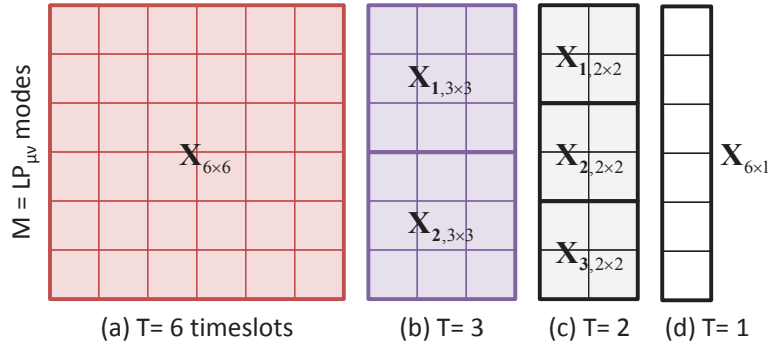


FIGURE E.14: Quatre configuration de mots de codes répartis en espace et en temps : (a) code 6×6 en un bloc unique, (b) deux blocs de codes 3×3 , (c) trois blocs de codes 2×2 , (d) Simple multiplexage spatial.

Pour illustrer les avantages de cette solution, nous considérons un système à 2 blocs où chacun est codé avec un code TAST 3×3 , et un système à 3 blocs où chacun est codé par le Silver code. La stratégie d'allocation des modes aux différents blocs va dépendre des disparités modales des gain dans les FMA en question. Les modes les plus atténués seront codés avec un mode privilégié pour moyenner au mieux les disparités de gain. Ainsi, pour la technologie FMA considérée dans notre étude, le schéma à 2 blocs sera composé des modes LP_{01} et $LP_{21a,b}$ dans un bloc, et des modes LP_{02} et $LP_{11a,b}$ dans l'autre ; tandis que le schéma à 3 blocs sera composé des blocs suivants : $\{LP_{01}, LP_{21a}\}$, $\{LP_{02}, LP_{21b}\}$ et $\{LP_{11a}, LP_{11b}\}$.

Pour ces solutions de codage multi-bloc, le modèle de canal vectorisé est le même que pour la solution à bloc ST unique dans (E.35) avec une matrice canal équivalente $\mathbf{H}_{eq}$ réarrangée, de plus petites dimensions. La matrice génératrice $\mathbf{M}_G$ représentant la configuration de codage ST multi-bloc sera composée d'une matrice diagonale en blocs avec les matrices génératrices des codes TAST 3×3 ou des codes Silver 2×2 sur sa diagonale, comme illustré ci-dessous :

$$\begin{aligned} \mathbf{M}_{G,2\text{-blocks}} &= \begin{bmatrix} \mathbf{M}_{G,\text{TAST}} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{G,\text{TAST}} \end{bmatrix} \text{ or} \\ \mathbf{M}_{G,3\text{-blocks}} &= \begin{bmatrix} \mathbf{M}_{G,\text{Silver}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{G,\text{Silver}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_{G,\text{Silver}} \end{bmatrix} \end{aligned} \quad (\text{E.39})$$

Nous évaluons les TEB moyens de ces solutions de codage en blocs en utilisant un décodeur ML ainsi que le même système de transmission MDM de 6 modes. La figure E.15 montre les courbes TEB pour trois taux de couplage différents, et avec des brouilleurs de mode après chaque FMA. Le schéma à 2 blocs absorbe toute la pénalité SNR à un $\text{TEB} = 10^{-3}$ pour les couplages moyen et fort tandis que la pénalité est réduite à 0.5dB pour du couplage faible. Le schéma à 3

blocs présente une pénalité SNR de 0.2dB dans les scénarios de couplage moyen et fort, et 1dB de pénalité avec du couplage faible. La pénalité accrue des schémas multi-blocs est due au fait que les codes utilisés moyennent les pertes observées à l'intérieur de chaque bloc, et combattent ainsi la MDL intra-bloc en négligeant la MDL inter-bloc, tandis que les systèmes où un bloc codé unique est appliqué moyenne les pertes sur tous les modes. Cependant, la complexité du décodage ML est considérablement réduite avec l'approche multi-bloc (décodage $\approx 10^{14}$ fois plus rapide). De plus, cette approche offre des solutions flexibles et évolutives qui peuvent être étendues à des systèmes MDM de plus grandes dimensions. Elle peut être également adapté à d'autres systèmes MDM où une technologie différente de FMA est utilisée ou d'autres sources de MDL existent.

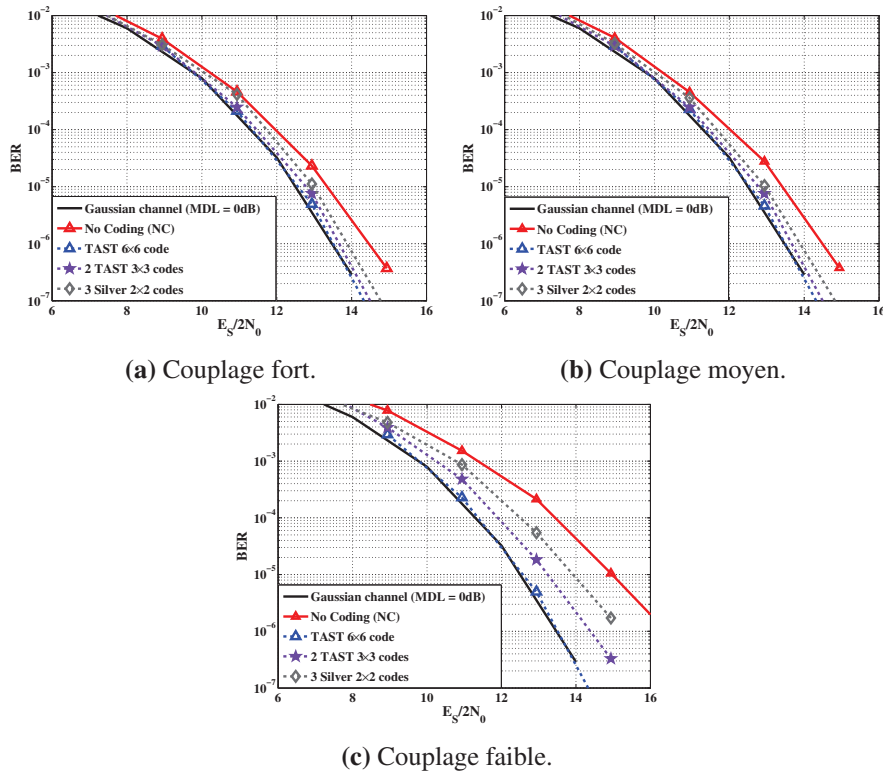


FIGURE E.15: TEB moyens en fonction du SNR des schémas de codage multi-bloc du système MDM à 6 modes avec brouillage de modes.

E.3.5 Conclusion

Dans cette section, nous avons montré, par le biais des simulations numériques, que le codage spatio-temporel (ST) est une technique prometteuse pour les systèmes MDM-OFDM impacté par la MDL. Cette technique peut être utilisée en tant que solution autonome ou en complément des solutions optiques telles que le brouillage de modes, en fonction du taux de couplage inter-modal dans les fibres FMF installées. La performance d'un décodage optimal des codes TAST a été étudiée sur des systèmes MDM de 3 et de 6 modes. Les pénalités engendrées par des niveaux de

MDL aussi élevés que 10dB ont pu être complètement éliminées. En outre, deux stratégies ont été présentées afin de réduire la complexité de décodage des schémas ST : d'une part, un décodeur de faible complexité a montré une performance quasi-optimale et d'autre part, une approche de codage multi-bloc a assuré des gains de codage intéressants et un grand potentiel de mise à l'échelle pour les systèmes MDM plus larges.

Les gains de codage obtenus peuvent assouplir les exigences en termes de MDL sur les composants optiques, ou d'une manière équivalente augmenter la portée des systèmes MDM. Ces observations ouvrent la voie à d'autres études intéressantes sur les solutions de codage ST pour les systèmes SDM telles que l'application d'autres familles de codes ST, ainsi que l'extension de l'étude au-delà de 6 modes (systèmes MDM avec un plus grand nombre de modes M ou systèmes MDM multiplexés en polarisation où la PDL peut être ajoutée comme autre facteur limitant).

E.4 Perspectives

Les conclusions des deux principales parties de cette thèse ouvrent la voie à d'autres études sur les avantages des solutions de codage ST à rendement plein pour les systèmes MIMO optiques. Parmi les nombreux axes de recherche possibles, nous développons quatre principaux axes que nous considérons comme les plus intéressants :

- La conception de codes pour les communications MIMO optiques : en se basant sur le critère dérivé pour les systèmes PDM, nous pouvons rechercher des schémas de codage qui répondent au mieux aux contraintes d'orthogonalité sans réduire le gain de multiplexage. En outre, la borne supérieure de probabilité d'erreur que nous avons trouvé et le critère de construction qui en dérive, peuvent être analysées plus en détail pour inclure les statistiques de la PDL comme sa distribution Maxwellienne. Par ailleurs, une analyse théorique peut aussi être effectuée pour les systèmes MDM impactés par la MDL pour en extraire les règles de construction de codes spécifiques à ces systèmes. Pour y arriver, la définition d'un modèle canal de forme compacte et relativement simple est nécessaire.
- Amélioration des modèles canal SDM et validation expérimentale : afin de nous concentrer sur la MDL, nous n'avons pas tenu compte des modes de polarisation ainsi que des solutions de multiplexage spatial hybrides MCF/MMF. Nous pouvons ainsi envisager de définir de nouveaux modèles canal SDM avec différentes sources de MDL et de PDL, et de tester de nouveaux schémas de codage ST sur ces canaux. Dans un second temps, des expériences de transmission SDM peuvent être effectuées pour tenir compte de l'interaction de la dispersion modale et de la MDL qui se traduira par une variation en fréquence de la MDL. L'expérience permettra de tenir compte aussi de la non-blanchité spatiale du bruit et, éventuellement, des effets non-linéaires pour des niveaux de puissance élevés injectés dans la fibre.

- Des solutions DSP à complexité réduite pour les communications MIMO optiques : nous avons choisi le décodeur par sphères parmi les multiples solutions de décodage en réseau de points pour implémenter le critère ML. D'autres décodeurs peuvent être testés comme le décodeur par pile (stack decoder) ou sa version de plus faible complexité : spherical-bounded stack decoder. Un décodage conjoint du code spatio-temporel et du codage FEC peut également être envisagée. En outre, l'approche de codage multi-bloc et la sélection de mode spatiaux peuvent être étendues à des solutions MCF/MMF hybrides où différents niveaux de MDL peuvent être observées parmi les sous-ensembles des canaux disponibles en fonction de la technologie des composants optiques SDM sous-jacents.
- La réduction des effets non-linéaires en utilisant du codage MIMO à rendement plein : un autre axe de recherche important dans le domaine des systèmes de transmission optique est la compensation des effets non-linéaires. L'évaluation de la capacité du codage MIMO à compenser ces effets dans divers systèmes MIMO optiques (PDM, SDM et WDM) demeure une piste intéressante.

Bibliography

- [1] D.J. Richardson, J.M. Fini, and L.E. Nelson. Space-division multiplexing in optical fibres. *Nature Photonics*, 7(5):354–362, 2013. Cited pages [xvii](#), [3](#), [9](#), [37](#), [152](#), [188](#), and [211](#)
- [2] A. Lobato, F. Ferreira, B. Inan, S. Adhikari, M. Kuschnerov, A. Napoli, B. Spinnler, and B. Lankl. Maximum-likelihood detection in few-mode fiber transmission with mode-dependent loss. *Photonics Technology Letters, IEEE*, 25(12):1095–1098, June 2013. Cited pages [xviii](#), [42](#), [43](#), [44](#), [123](#), [124](#), [137](#), [212](#), [213](#), and [214](#)
- [3] T. Mizuochi. Recent progress in forward error correction and its interplay with transmission impairments. *Selected Topics in Quantum Electronics, IEEE Journal of*, 12(4):544–554, July 2006. Cited pages [xviii](#), [52](#), and [53](#)
- [4] CISCO. The zettabyte era—trends and analysis. *Cisco Public Information*, 2014. Cited pages [1](#) and [185](#)
- [5] P.P. Mitra and J.B. Stark. Nonlinear limits to the information capacity of optical fibre communications. *Nature*, 411:1027–1030, 2001. Cited pages [3](#), [8](#), and [188](#)
- [6] R. Essiambre, G. Kramer, P.J. Winzer, G.J. Foschini, and B. Goebel. Capacity limits of optical fiber networks. *Lightwave Technology, Journal of*, 28(4):662–701, Feb 2010. Cited pages [3](#), [8](#), [49](#), [68](#), [69](#), and [188](#)
- [7] P. Bayvel, C. behrens, and D.S. Millar. Dsp and its applications in optical communications systems, 2013. Cited pages [3](#) and [188](#)
- [8] S. Mumtaz, G. Rekaya-Ben, and Y. Jaouën. Space-time codes for optical fiber communication with polarization multiplexing. *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5, May 2010. Cited pages [4](#), [50](#), [86](#), [88](#), [89](#), [94](#), [102](#), [189](#), and [192](#)
- [9] Pierre Sillard, Marianne Astruc, David Boivin, Hélène Maerten, and Lionel Provost. Few-mode fiber for uncoupled mode-division multiplexing transmissions. *37th European Conference and Exposition on Optical Communications*, page Tu.5.LeCervin.7, 2011. Cited pages [5](#), [37](#), [41](#), and [190](#)
- [10] Lars Grüner-Nielsen, Yi Sun, Jeffrey W. Nicholson, Dan Jakobsen, Kim G. Jespersen, Jr. Robert Lingle, and Bera Pálsdóttir. Few mode transmission fiber with low dgd, low mode coupling, and low loss. *J. Lightwave Technol.*, 30(23):3693–3698, Dec 2012. Cited pages [5](#), [9](#), [37](#), [41](#), [133](#), [190](#), [211](#), and [212](#)
- [11] J. Proakis and M. Salehi. *Digital Communications, Fifth Edition*. Mc Graw - Hill International Edition, 2008. Cited pages [8](#), [51](#), [58](#), [72](#), [73](#), [81](#), [85](#), [108](#), [110](#), [200](#), and [202](#)
- [12] P.J. Winzer. Spatial multiplexing in fiber optics: The 10x scaling of metro/core capacities. *Bell Labs Technical Journal*, 19:22–30, 2014. Cited page [8](#)
- [13] Alan H. Gnauck, Sethumadhavan Chandrasekhar, Xiang Liu, Sebastian Randel, Steve Corteselli, Thierry Taunay, Benyuan Zhu, and Mikhail Fishteyn. Wdm transmission of 603-gb/s superchannels over 845 km of 7-core fiber with 42.2 b/s/hz spectral efficiency. *European Conference and Exhibition on Optical Communication*, page Th.2.C.2, 2012. Cited pages [9](#), [37](#), and [211](#)

BIBLIOGRAPHY

- [14] Charles Cox, Edward Ackerman, Roger Helkey, and Gary E Betts. Techniques and performance of intensity-modulation direct-detection analog optical links. *IEEE Transactions on Microwave theory and techniques*, 45(8):1375–1383, 1997. Cited page [11](#)
- [15] S. Chandrasekhar, Xiang Liu, P.J. Winzer, J.E. Simsarian, and R.A Griffin. Compact all-inp laser-vector-modulator for generation and transmission of 100-gb/s pdm-qpsk and 200-gb/s pdm-16-qam. *Lightwave Technology, Journal of*, 32(4):736–742, Feb 2014. Cited pages [15](#) and [32](#)
- [16] Mengdi Song, Erwan Pincemin, Didier Grot, Thierry Guillossou, Yves Jaouën, and Raphaël Le Bidan. Robustness of coherent 100 gbps qpsk and 200 gbps 16qam-ofdm to practical implementation impairments. In *Advanced Photonics for Communications*, page SM3E.5. Optical Society of America, 2014. Cited pages
- [17] Dayou Qian, Ming-Fang Huang, E. Ip, Yue-Kai Huang, Yin Shao, Junqiang Hu, and Ting Wang. 101.7-tb/s (370x294-gb/s) pdm-128qam-ofdm transmission over 3x55-km ssmf using pilot-based phase noise mitigation. *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011. Cited page [15](#)
- [18] H.-G. Unger. *Planar Optical Waveguides and Fibers*. Oxford, 1977. Cited pages [15](#), [16](#), [38](#), [39](#), [40](#), and [123](#)
- [19] G. Agrawal. *Non-Linear Fiber Optics, Fourth Edition*. Academic Press, 2007. Cited pages [16](#), [17](#), [21](#), [26](#), [30](#), [31](#), [38](#), and [101](#)
- [20] E. Desurvire, J. R. Simpson, and P. C. Becker. High-gain erbium-doped traveling-wave fiber amplifier. *Opt. Lett.*, 12(11):888–890, Nov 1987. Cited page [18](#)
- [21] R.J. Mears, L. Reekie, I.M. Jauncey, and D.N. Payne. Low-noise erbium-doped fibre amplifier operating at 1.54 μ m. *Electronics Letters*, 23:1026–1028(2), September 1987. Cited page [18](#)
- [22] Philippe M Becker, Anders A Olsson, and Jay R Simpson. *Erbium-doped fiber amplifiers: fundamentals and technology*. Academic press, 1999. Cited pages [19](#) and [21](#)
- [23] J. Hegarty, N.A. Olsson, and L. Goldner. Cw pumped raman preamplifier in a 45 km-long fibre transmission system operating at 1.5 μ m and 1 gbit/s. *Electronics Letters*, 21:290–292(2), March 1985. Cited page [20](#)
- [24] Yann Frignac, Jean-Christophe Antona, and Sébastien Bigo. Enhanced analytical engineering rule for fast optimization of dispersion maps in 40 gbit/s-based transmission systems. *Optical Fiber Communication Conference*, page TuN3, 2004. Cited page [21](#)
- [25] Seb J. Savory. Digital filters for coherent optical receivers. *Optics Express*, 16(2):804–817, 2008. Cited pages [21](#), [25](#), [28](#), [32](#), [54](#), [56](#), [191](#), and [192](#)
- [26] S. V. Chernikov and J. R. Taylor. Measurement of normalization factor of n_2 for random polarization in optical fibers. *Opt. Lett.*, 21(19):1559–1561, Oct 1996. Cited page [21](#)
- [27] Keang-Po Ho. The effect of interferometer phase error on direct-detection dpsk and dqpsk signals. *Photonics Technology Letters, IEEE*, 16(1):308–310, Jan 2004. Cited page [23](#)
- [28] Kazuro Kikuchi. Coherent optical communication systems. *Optical Fiber Telecommunications VB: systems and networks*, 2008. Cited page [24](#)
- [29] Gabriel Charlet. Coherent detection associated with digital signal processing for fiber optics communication. *Comptes Rendus Physique*, 9(9–10):1012 – 1030, 2008. Recent advances in optical telecommunications / Avancés récentes en télécommunications optiques. Cited page [24](#)

-
- [30] R.A Linke and A.H. Gnauck. High-capacity coherent lightwave systems. *Lightwave Technology, Journal of*, 6(11):1750–1769, Nov 1988. Cited page [24](#)
 - [31] R.I Killey, P.M. Watts, M. Glick, and P. Bayvel. Electronic precompensation techniques to combat dispersion and nonlinearities in optical transmission. *Optical Communication, 2005. ECOC 2005. 31st European Conference on*, 2:251–254 vol.2, Sept 2005. Cited page [24](#)
 - [32] J.M. Gene, P.J. Winzer, S. Chandrasekhar, and H. Kogelnik. Joint pmd and chromatic dispersion compensation using an mlse. *Optical Communications, 2006. ECOC 2006. European Conference on*, pages 1–2, Sept 2006. Cited pages [24](#) and [31](#)
 - [33] A.J. Viterbi and A.M. Viterbi. Nonlinear estimation of psk-modulated carrier phase with application to burst digital transmission. *Information Theory, IEEE Transactions on*, 29(4):543–551, Jul 1983. Cited page [25](#)
 - [34] F. Forghieri, R.W. Tkach, A.R. Chraplyvy, and D. Marcuse. Reduction of four-wave mixing crosstalk in wdm systems using unequally spaced channels. *Photonics Technology Letters, IEEE*, 6(6):754–756, June 1994. Cited page [26](#)
 - [35] A.R. Chraplyvy, A.H. Gnauck, R.W. Tkach, and R. M. Derosier. 8*10 gb/s transmission through 280 km of dispersion-managed fiber. *Photonics Technology Letters, IEEE*, 5(10):1233–1235, Oct 1993. Cited page [26](#)
 - [36] Peter J. Winzer and René-Jean Essiambre. Advanced modulation formats for high-capacity optical transport networks. *J. Lightwave Technol.*, 24(12):4711–4728, Dec 2006. Cited page [27](#)
 - [37] Marco Secondini. Receiver training for efficient nonlinear equalization and detection in optical communications. *Advanced Photonics for Communications*, page ST2D.4, 2014. Cited page [27](#)
 - [38] [100 GE products by Alcatel-Lucent, Fujitsu, and NEC](#). [Online; accessed 9-August-2014]. Cited page [27](#)
 - [39] E. Agrell and M. Karlsson. Power-efficient modulation formats in coherent transmission systems. *Lightwave Technology, Journal of*, 27(22):5115–5126, Nov 2009. Cited pages [27](#), [54](#), and [87](#)
 - [40] D. van den Borne, S. L. Jansen, E. Gottwald, P. M. Krummrich, G. D. Khoe, and H. de Waardt. 1.6-b/s/hz spectrally efficient transmission over 1700 km of ssmf using 40x85.6-gb/s polmux-rz-dqpsk. *J. Lightwave Technol.*, 25(1):222–232, Jan 2007. Cited page [28](#)
 - [41] C.R.S. Fludger, T. Duthel, Dirk Van den Borne, C. Schulien, E. D Schmidt, T. Wuth, J. Geyer, E. De Man, G-D Khoe, and H. de Waardt. Coherent equalization and polmux-rz-dqpsk for robust 100-ge transmission. *Lightwave Technology, Journal of*, 26(1):64–72, Jan 2008. Cited page [28](#)
 - [42] Sander L. Jansen, Itsuro Morita, and Hideaki Tanaka. 16x52.5-gb/s, 50-ghz spaced, polmux-co-ofdm transmission over 4,160 km of ssmf enabled by mimo processing. *Optical Communication - Post-Deadline Papers (published 2008), 2007 33rd European Conference and Exhibition of*, pages 1–2, Sept 2007. Cited pages [28](#), [32](#), [62](#), [69](#), [191](#), and [194](#)
 - [43] Sander Lars Jansen, Itsuro Morita, TimC. W. Schenk, and Hideaki Tanaka. 121.9-gb/s pdm-ofdm transmission with 2-b/s/hz spectral efficiency over 1000 km of ssmf. *J. Lightwave Technol.*, 27(3):177–188, Feb 2009. Cited pages [28](#), [32](#), [62](#), [69](#), [191](#), and [194](#)
 - [44] Ezra Ip, Alan Pak Tao Lau, Daniel J. F. Barros, and Joseph M. Kahn. Coherent detection in optical fiber systems. *Opt. Express*, 16(2):753–791, Jan 2008. Cited page [29](#)
 - [45] H. Bülow, W. Baumert, H. Schmuck, F. Mohr, T. Schulz, F. Küppers, and W. Weiershausen. Measurement of the maximum speed of pmd fluctuation in installed field fiber. *Optical Fiber Communication Conference and the International Conference on Integrated Optics and Optical Fiber Communication*, page WE4, 1999. Cited pages [30](#), [195](#), and [197](#)
-

- [46] L. E. Nelson and R. M. Jopson. Introduction to polarization mode dispersion in optical systems. *Journal of Optical and Fiber Communications Reports*, 1(4):312–344, 2004. Cited pages [30](#), [31](#), and [195](#)
- [47] F. Heismann, D.A Fishman, and D.L. Wilson. Automatic compensation of first order polarization mode dispersion in a 10 gb/s transmission system. *Optical Communication, 1998. 24th European Conference on*, 1:529–530 vol.1, Sep 1998. Cited page [31](#)
- [48] G.D. Forney. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *Information Theory, IEEE Transactions on*, 18(3):363–378, May 1972. Cited page [31](#)
- [49] Akihide Sano, Hiroji Masuda, Takayuki Kobayashi, Masamichi Fujiwara, Kengo Horikoshi, Eiji Yoshida, Yutaka Miyamoto, Munehiro Matsui, Masato Mizoguchi, Hiroshi Yamazaki, Yohei Sakamaki, and Hiroyuki Ishii. 69.1-tb/s (432 x 171-gb/s) c- and extended l-band transmission over 240 km using pdm-16-qam modulation and digital coherent detection. *Optical Fiber Communication Conference*, page PDPB7, 2010. Cited page [32](#)
- [50] Thomas Duthel, Chris R. Fludger, Jonas Geyer, and Christoph Schulien. Impact of polarisation dependent loss on coherent polmux-nrz-dqpsk. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference*, page OThU5, 2008. Cited pages [33](#), [34](#), [191](#), and [196](#)
- [51] Francisco A. Garcia, Darli A. Mello, and Helio Waldman. Feedforward carrier recovery for polarization demultiplexed signals with unequal signal to noise ratios. *Opt. Express*, 17(10):7958–7969, May 2009. Cited page [33](#)
- [52] Lynn E. Nelson, Cristian Antonelli, Antonio Mecozzi, Martin Birk, Peter Magill, Anton Schex, and Lutz Rapp. Statistics of polarization dependent loss in an installed long-haul wdm system. *Opt. Express*, 19(7):6790–6796, Mar 2011. Cited pages [33](#), [35](#), [196](#), and [210](#)
- [53] A Mecozzi and M. Shtaif. The statistics of polarization-dependent loss in optical communication systems. *Photonics Technology Letters, IEEE*, 14(3):313–315, March 2002. Cited pages [33](#), [96](#), [117](#), [196](#), and [209](#)
- [54] N Gisin. Statistics of polarization dependent losses. *Optics Communications*, 114(5–6):399 – 405, 1995. Cited pages [33](#) and [196](#)
- [55] Kunihiro Mori, Tomoyoshi Kataoka, Takayuki Kobayashi, and Shingo Kawai. Statistics and performance under combined impairments induced by polarization-dependent-loss in polarization-division-multiplexing digital coherent transmission. *Opt. Express*, 19(26):B673–B680, Dec 2011. Cited page [34](#)
- [56] Lynn E. Nelson, Guodong Zhang, Martin Birk, Craig Skolnick, Rejoy Isaac, Y. Pan, Christian Rasmussen, Graeme Pendock, and Benny Mikkelsen. A robust real-time 100g transceiver with soft-decision forward error correction. *J. Opt. Commun. Netw.*, 4(11):B131–B141, Nov 2012. Cited pages [34](#), [51](#), and [53](#)
- [57] B. Huttner, C. Geiser, and N. Gisin. Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses. *Selected Topics in Quantum Electronics, IEEE Journal of*, 6(2):317–329, March 2000. Cited page [34](#)
- [58] A. O. Lima, Jr. I. T. Lima, C. R. Menyuk, and T. Adali. Comparison of penalties resulting from first-order and all-order polarization mode dispersion distortions in optical fiber transmission systems. *Opt. Lett.*, 28(5):310–312, Mar 2003. Cited pages [34](#), [97](#), and [196](#)
- [59] Mark Shtaif. Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss. *Opt. Express*, 16(18):13918–13932, Sep 2008. Cited pages [35](#), [99](#), [103](#), and [197](#)
- [60] Chongjin Xie. Polarization-dependent loss induced penalties in pdm-qpsk coherent optical communication systems. *Optical Fiber Communication Conference*, page OWE6, 2010. Cited pages [35](#), [99](#), [111](#), and [207](#)

-
- [61] Olga Vassilieva, Inwoong Kim, Youichi Akasaka, Martin Bouda, and Motoyoshi Sekiya. Interplay between pdl and nonlinear effects in coherent polarization multiplexed systems. *Opt. Express*, 19(26):B357–B362, Dec 2011. Cited pages [35](#), [111](#), [196](#), and [207](#)
 - [62] Mahdi Zamani, Zhuhong Zhang, Chen Chen, and Chuandong Li. Pdl compensation using whitening matrix in polarization division multiplexed coherent optical transmission. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, page JW2A.49, 2013. Cited page [35](#)
 - [63] Nelson J. Muga and Armando Nolasco Pinto. Digital pdl compensation in 3d stokes space. *J. Lightwave Technol.*, 31(13):2122–2130, Jul 2013. Cited page [35](#)
 - [64] Peter J. Winzer and Gerard J. Foschini. MIMO capacities and outage probabilities in spatially multiplexed optical transport systems. *Opt. Express*, 19(17):16680–16696, Aug 2011. Cited pages [36](#), [39](#), [42](#), [124](#), [125](#), [127](#), [132](#), [145](#), [161](#), and [212](#)
 - [65] P.J. Winzer. Energy-efficient optical transport capacity scaling through spatial multiplexing. *Photonics Technology Letters, IEEE*, 23(13):851–853, July 2011. Cited page [36](#)
 - [66] K. Takenaga, Y. Sasaki, Ning Guan, S. Matsuo, M. Kasahara, K. Saitoh, and M. Koshiba. Large effective-area few-mode multicore fiber. *Photonics Technology Letters, IEEE*, 24(21):1941–1944, Nov 2012. Cited pages [37](#) and [211](#)
 - [67] R. Ryf, N.K. Fontaine, and R. Essiambre. Spot-based mode coupler for mode-multiplexed transmission in few-mode fiber. *Photonics Society Summer Topical Meeting Series, 2012 IEEE*, pages 199–200, July 2012. Cited page [37](#)
 - [68] Peter M. Krummrich. Optical amplification and optical filter based signal processing for cost and energy efficient spatial multiplexing. *Opt. Express*, 19(17):16636–16652, Aug 2011. Cited pages [37](#), [42](#), [124](#), [211](#), and [212](#)
 - [69] S. Iano, T. Sato, S. Sentsui, T. Kuroha, and Y. Nishimura. Multicore optical fiber. *Optical Fiber Communication*, page WB1, 1979. Cited page [37](#)
 - [70] Jun Sakaguchi, Benjamin J. Puttnam, Werner Klaus, Yoshinari Awaji, Naoya Wada, Atsushi Kanno, Tetsuya Kawanishi, Katsunori Imamura, Harumi Inaba, Kazunori Mukasa, Ryuichi Sugizaki, Tetsuya Kobayashi, and Masayuki Watanabe. 19-core fiber transmission of 19x100x172-gb/s sdm-wdm-pdm-qpsk signals at 305tb/s. *National Fiber Optic Engineers Conference*, page PDP5C.1, 2012. Cited page [37](#)
 - [71] Hidehiko Takara, Akihide Sano, Takayuki Kobayashi, Hirokazu Kubota, Hiroto Kawakami, Akihiko Matsuura, Yutaka Miyamoto, Yoshiteru Abe, Hirotaka Ono, Kota Shikama, Yukihiro Goto, Kyoza Tsujikawa, Yusuke Sasaki, Itaru Ishida, Katsuhiko Takenaga, Shoichiro Matsuo, Kunimasa Saitoh, Masanori Koshiba, and Toshio Morioka. 1.01-pb/s (12 sdm/222 wdm/456 gb/s) crosstalk-managed transmission with 91.4-b/s/hz aggregate spectral efficiency. *European Conference and Exhibition on Optical Communication*, page Th.3.C.1, 2012. Cited page [38](#)
 - [72] B. Zhu, T. F. Taunay, M. F. Yan, J. M. Fini, M. Fishteyn, E. M. Monberg, and F. V. Dimarcello. Seven-core multicore fiber transmissions for passive optical network. *Opt. Express*, 18(11):11117–11122, May 2010. Cited page [38](#)
 - [73] R. Ryf, R. Essiambre, S. Randel, A.H. Gnauck, P.J. Winzer, T. Hayashi, T. Taru, and T. Sasaki. MIMO-based crosstalk suppression in spatially multiplexed 3x56-gb/s pdm-qpsk signals for strongly coupled three-core fiber. *Photonics Technology Letters, IEEE*, 23(20):1469–1471, Oct 2011. Cited page [38](#)
 - [74] Keang-Po Ho and Joseph M. Kahn. Mode-dependent loss and gain: statistics and effect on mode-division multiplexing. *Opt. Express*, 19(17):16612–16635, Aug 2011. Cited pages [39](#), [40](#), [42](#), [43](#), [122](#), [123](#), [124](#), [125](#), [135](#), [136](#), [145](#), [212](#), [214](#), [216](#), [217](#), and [218](#)
-

- [75] KP Ho and JM Kahn. Mode coupling and its impact on spatially multiplexed systems, 2013. Cited page [39](#)
- [76] F. Ferreira, S. Jansen, P. Monteiro, and H. Silva. Nonlinear semi-analytical model for simulation of few-mode fiber transmission. *Photonics Technology Letters, IEEE*, 24(4):240–242, Feb 2012. Cited pages [40](#), [133](#), [214](#), and [215](#)
- [77] F. Yaman, E. Mateo, and Ting Wang. Impact of modal crosstalk and multi-path interference on few-mode fiber transmission. *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012. Cited pages [40](#), [133](#), and [214](#)
- [78] A.A. Juarez, E. Krune, S. Warm, C.A. Bunge, and K. Petermann. Modeling of mode coupling in multimode fibers with respect to bandwidth and loss. *Lightwave Technology, Journal of*, 32(8):1549–1558, April 2014. Cited pages [40](#), [42](#), [123](#), [124](#), [128](#), [133](#), [214](#), and [216](#)
- [79] Clemens Koebele, Massimiliano Salsi, Laurent Milord, Roland Ryf, Cristian A. Bolle, Pierre Sillard, Sebastien Bigo, and Gabriel Charlet. 40km transmission of five mode division multiplexed data streams at 100gb/s with low mimo-dsp complexity. *37th European Conference and Exposition on Optical Communications*, page Th.13.C.3, 2011. Cited pages [40](#) and [214](#)
- [80] An Li, Abdullah Al Amin, Xi Chen, and William Shieh. Transmission of 107-gb/s mode and polarization multiplexed co-ofdm signal over a two-mode fiber. *Opt. Express*, 19(9):8808–8814, Apr 2011. Cited pages [40](#), [145](#), and [214](#)
- [81] P.J. Winzer, AH. Gnauck, A Konczykowska, F. Jorge, and J-Y Dupuy. Penalties from in-band crosstalk for advanced optical modulation formats. *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, Sept 2011. Cited page [40](#)
- [82] Sebastian Randel, Roland Ryf, Alan Gnauck, Miguel A. Mestre, Christian Schmidt, Rene Essiambre, Peter Winzer, Roger Delbue, Peter Pupalaikis, Anirudh Sureka, Yi Sun, Xinli Jiang, and Robert Lingle. Mode-multiplexed 6x20-gbd qpsk transmission over 1200-km dgd-compensated few-mode fiber. *National Fiber Optic Engineers Conference*, page PDP5C.5, 2012. Cited pages [40](#), [41](#), [122](#), [211](#), and [212](#)
- [83] R. Ryf, S. Randel, A.H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaelpour, E.C. Burrows, R. Essiambre, P.J. Winzer, D.W. Peckham, A.H. McCurdy, and R. Lingle. Mode-division multiplexing over 96 km of few-mode fiber using coherent 6x6 mimo processing. *Lightwave Technology, Journal of*, 30(4):521–531, Feb 2012. Cited pages [40](#), [41](#), [122](#), [133](#), [212](#), and [214](#)
- [84] S.O. Arik, D. Askarov, and J.M. Kahn. Effect of mode coupling on signal processing complexity in mode-division multiplexing. *Lightwave Technology, Journal of*, 31(3):423–431, Feb 2013. Cited pages [40](#), [42](#), [43](#), [125](#), [212](#), and [214](#)
- [85] Sebastian Randel, Roland Ryf, Alberto Sierra, Peter J. Winzer, Alan H. Gnauck, Cristian A. Bolle, René-Jean Essiambre, David W. Peckham, Alan McCurdy, and Robert Lingle. 6x56-gb/s mode-division multiplexed transmission over 33-km few-mode fiber enabled by 6x6 mimo equalization. *Opt. Express*, 19(17):16697–16707, Aug 2011. Cited page [40](#)
- [86] Beril Inan, Yongmin Jung, Vincent Sleiffer, Maxim Kuschnerov, Lars Gruner-Nielsen, Susmita Adhikari, Sander L. Jansen, David J. Richardson, Shaif ul Alam, Bernhard Spinnler, and Norbert Hanik. Low computational complexity mode division multiplexed ofdm transmission over 130 km of few mode fiber. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, page OW4F.4, 2013. Cited pages [41](#) and [151](#)

-
- [87] Takayoshi Mori, Taiji Sakamoto, Masaki Wada, Takashi Yamamoto, and Fumihiko Yamamoto. Few-mode fibers supporting more than two lp modes for mode-division-multiplexed transmission with mimo dsp. *J. Lightwave Technol.*, 32(14):2468–2479, Jul 2014. Cited page [41](#)
 - [88] Vincent Sleiffer, Yongmin Jung, Beril Inan, Haoshuo Chen, Roy van Uden, Maxim Kuschnerov, Dirk van den Borne, Sander Jansen, Vladimir Veljanovski, Ton Koonen, David Richardson, Shaif ul Alam, Francesco Poletti, Jayanta Sahu, Anirban Dhar, Brian Corbett, Richard Winfield, Andrew Ellis, and Huug De Waardt. Mode-division-multiplexed 3x112-gb/s dp-qpsk transmission over 80 km few-mode fiber with inline mm-edfa and blind dsp. *European Conference and Exhibition on Optical Communication*, page Tu.1.C.2, 2012. Cited page [41](#)
 - [89] Roland Ryf, Nicolas K. Fontaine, Miquel A. Mestre, Sebastian Randel, Xavi Palou, Cristian Bolle, Alan H. Gnauck, Sethumadhavan Chandrasekhar, Xiang Liu, Binbin Guan, René-Jean Essiambre, Peter J. Winzer, Sergio Leon-Saval, Joss Bland-Hawthorn, Roger Delbue, Peter Pupaikis, Anirudh Sureka, Yi Sun, Lars Grüner-Nielsen, Rasmus V. Jensen, and Robert Lingle. 12 x 12 mimo transmission over 130-km few-mode fiber. *Frontiers in Optics 2012/Laser Science XXVIII*, page FW6C.4, 2012. Cited pages [42](#) and [211](#)
 - [90] Stefan Warm and Klaus Petermann. Splice loss requirements in multi-mode fiber mode-division-multiplex transmission links. *Opt. Express*, 21(1):519–532, Jan 2013. Cited pages [42](#), [43](#), [44](#), [123](#), [124](#), [127](#), [128](#), [130](#), [132](#), [133](#), [134](#), [212](#), [214](#), [215](#), [216](#), [217](#), and [222](#)
 - [91] N. Cvijetic, E. Ip, N. Prasad, and Ming-Jun Li. Experimental frequency-domain channel matrix characterization for sdm-mimo-ofdm systems. *Photonics Society Summer Topical Meeting Series, 2013 IEEE*, pages 139–140, July 2013. Cited pages [42](#) and [135](#)
 - [92] Guillaume Le Cocq, Yves Quinquempois, Antoine Le Rouge, Géraud Bouwmans, Hicham El Hamzaoui, Karen Delplace, Mohamed Bouazaoui, and Laurent Bigot. Few mode er3+ doped fiber with micro-structured core for mode division multiplexing in the c-band. *Opt. Express*, 21(25):31646–31659, Dec 2013. Cited pages [42](#), [122](#), [124](#), [133](#), [135](#), [211](#), [212](#), [213](#), [215](#), and [217](#)
 - [93] Joel Carpenter, Benjamin J. Eggleton, and Jochen Schröder. 110x110 optical mode transfer matrix inversion. *Opt. Express*, 22(1):96–101, Jan 2014. Cited page [42](#)
 - [94] A. Lobato, F. Ferreira, J. Rabe, M. Kuschnerov, B. Spinnler, and B. Lankl. Mode scramblers and reduced-search maximum-likelihood detection for mode-dependent-loss-impaired transmission. *Optical Communication (ECOC 2013), 39th European Conference and Exhibition on*, pages 1–3, Sept 2013. Cited pages [43](#), [44](#), [130](#), [137](#), [140](#), [212](#), and [213](#)
 - [95] S. Mumtaz, R. Essiambre, and G.P. Agrawal. Nonlinear propagation in multimode and multicore fibers: Generalization of the manakov equations. *Lightwave Technology, Journal of*, 31(3):398–406, Feb 2013. Cited page [44](#)
 - [96] G.P. Agrawal and R.-J. Essiambre. Nonlinear limits of sdm transmission. *Photonics Society Summer Topical Meeting Series, 2014 IEEE*, pages 174–175, July 2014. Cited page [44](#)
 - [97] G. Rademacher, S. Warm, and K. Petermann. Influence of discrete mode coupling on the nonlinear interaction in mode-multiplexed systems. *Photonics Technology Letters, IEEE*, 25(13):1203–1206, July 2013. Cited page [44](#)
 - [98] Bruno Lavigne and Oriol Bertran-Pardo. 400gb/s real-time trials on commercial systems using bit-rate-adaptive transponders for next generation networks. *Advanced Photonics for Communications*, page SM3E.6, 2014. Cited page [44](#)
-

BIBLIOGRAPHY

- [99] S. Chandrasekhar and Xiang Liu. Experimental investigation on the performance of closely spaced multi-carrier pdm-qpsk with digital coherent detection. *Opt. Express*, 17(24):21350–21361, Nov 2009. Cited page [45](#)
- [100] G. Bosco, A. Carena, V. Curri, P. Poggiolini, and F. Forghieri. Performance limits of nyquist-wdm and co-ofdm in high-speed pm-qpsk systems. *Photonics Technology Letters, IEEE*, 22(15):1129–1131, Aug 2010. Cited page [45](#)
- [101] Xiang Liu, Sethumadhavan Chandrasekhar, Peter Winzer, Thorsten Lotz, John Carlson, Jenny Yang, Geoffrey Cheren, and Scott Zederbaum. 1.5-tb/s guard-banded superchannel transmission over 56x 100-km (5600-km) ulaf using 30-gbaud pilot-free ofdm-16qam signals with 5.75-b/s/hz net spectral efficiency. *European Conference and Exhibition on Optical Communication*, page Th.3.C.5, 2012. Cited page [45](#)
- [102] Lukas Kull, T. Toifl, M. Schmatz, P. A. Francese, C. Menolfi, M. Braendli, M. Kossel, T. Morf, T. Meyer Anderson, and Yusuf Leblebici. A 90gs/s 8b 667mw 64x interleaved sar adc in 32nm digital soi cmos. *Proceedings of the 2014 International Solid-State Circuits Conference (ISSCC)*, 2014. Cited pages [49](#), [51](#), and [152](#)
- [103] W. Shieh, X. Yi, Y. Ma, and Y. Tang. Theoretical and experimental study on pm-d supported transmission using polarization diversity in coherent optical ofdm systems. *Opt. Express*, 15(16):9936–9947, Aug 2007. Cited pages [50](#), [60](#), [61](#), [69](#), [191](#), [192](#), [193](#), and [194](#)
- [104] Beril Inan, Bernhard Spinnler, Filipe Ferreira, Dirk van den Borne, Adriana Lobato, Susmita Adhikari, Vincent A. J. M. Sleiffer, Maxim Kuschnerov, Norbert Hanik, and Sander L. Jansen. Dsp complexity of mode-division multiplexed receivers. *Opt. Express*, 20(10):10859–10869, May 2012. Cited pages [50](#), [57](#), [69](#), and [127](#)
- [105] C.E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, Jan 1949. Cited page [51](#)
- [106] T. Tsuritani, K. Ishida, A. Agata, K. Shimomura, I. Morita, T. Tokura, H. Taga, T. Mizuochi, N. Edagawa, and S. Akiba. 70-ghz-spaced 40 times;42.7 gb/s transpacific transmission over 9400 km using prefiltered csrz-dpsk signals, all-raman repeaters, and symmetrically dispersion-managed fiber spans. *Lightwave Technology, Journal of*, 22(1):215–224, Jan 2004. Cited page [53](#)
- [107] R.G. Gallager. Low-density parity-check codes. *Information Theory, IRE Transactions on*, 8(1):21–28, January 1962. Cited page [53](#)
- [108] IB. Djordjevic, M. Cvijetic, Lei Xu, and Ting Wang. Using ldpc-coded modulation and coherent detection for ultra highspeed optical transmission. *Lightwave Technology, Journal of*, 25(11):3619–3625, Nov 2007. Cited page [53](#)
- [109] IB. Djordjevic, B. Vasic, and M.A Neifeld. Ldpc-coded ofdm for optical communication systems with direct detection. *Selected Topics in Quantum Electronics, IEEE Journal of*, 13(5):1446–1454, Sept 2007. Cited page [53](#)
- [110] Ivan B. Djordjevic. Advanced coding for optical communications, 2013. Cited page [53](#)
- [111] T. Mizuochi, Y. Konishi, Y. Miyata, T. Inoue, K. Onohara, S. Kametani, T. Sugihara, K. Kubo, H. Yoshida, T. Kobayashi, and T. Ichikawa. Experimental demonstration of concatenated ldpc and rs codes by fpgas emulation. *Photonics Technology Letters, IEEE*, 21(18):1302–1304, Sept 2009. Cited page [53](#)
- [112] Sami Mumtaz, Ghaya Rekaya-Ben Othman, and Yves Jaouen. Quasi-cyclic ldpc based on peg construction for optical communications. *Advanced Photonics*, page SPWB2, 2011. Cited page [53](#)

-
- [113] Hongbin Zhang, Hussam G. Batshon, Dmitri Foursa, Matt Mazurczyk, Jin-Xing Cai, Carl Davidson, Alexei Pilipetskii, Georg Mohs, and Neal Bergano. 30.58 tb/s transmission over 7,230 km using pdm half 4d-16qam coded modulation with 6.1 b/s/hz spectral efficiency. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, page OTu2B.3, 2013. Cited pages [53](#) and [87](#)
 - [114] D. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *Communications, IEEE Transactions on*, 28(11):1867–1875, Nov 1980. Cited page [56](#)
 - [115] Charles Laperle, Bernard Villeneuve, Zhuhong Zhang, Doug McGhan, Han Sun, and Maurice O’Sullivan. Wdm performance and pm� tolerance of a coherent 40-gbit/s dual-polarization qpsk transceiver. *J. Lightwave Technol.*, 26(1):168–175, Jan 2008. Cited page [56](#)
 - [116] I Fatadin, D. Ives, and S.J. Savory. Blind equalization and carrier phase recovery in a 16-qam optical coherent system. *Lightwave Technology, Journal of*, 27(15):3042–3049, Aug 2009. Cited pages [56](#) and [58](#)
 - [117] Neng Bai and Guifang Li. Adaptive frequency-domain equalization for mode-division multiplexed transmission. *Photonics Technology Letters, IEEE*, 24(21):1918–1921, Nov 2012. Cited page [57](#)
 - [118] S. Randel, P.J. Winzer, M. Montoliu, and R. Ryf. Complexity analysis of adaptive frequency-domain equalization for mimo-sdm transmission. *Optical Communication (ECOC 2013), 39th European Conference and Exhibition on*, pages 1–3, Sept 2013. Cited page [57](#)
 - [119] Andreas Leven, N. Kaneda, Ut-Va Koc, and Young-Kai Chen. Frequency estimation in intradyne reception. *Photonics Technology Letters, IEEE*, 19(6):366–368, March 2007. Cited page [58](#)
 - [120] G.L. Stuber, J.R. Barry, S.W. McLaughlin, Ye Li, M.-A Ingram, and T.G. Pratt. Broadband mimo-ofdm wireless communications. *Proceedings of the IEEE*, 92(2):271–294, Feb 2004. Cited pages [59](#), [60](#), [64](#), [65](#), and [66](#)
 - [121] Jean Armstrong. Ofdm for optical communications. *J. Lightwave Technol.*, 27(3):189–204, Feb 2009. Cited pages [60](#) and [63](#)
 - [122] Qi Yang, Yan Tang, Y. Ma, and William Shieh. Experimental demonstration and numerical simulation of 107-gb/s high spectral efficiency coherent optical ofdm. *Lightwave Technology, Journal of*, 27(3):168–176, Feb 2009. Cited page [61](#)
 - [123] N. Cvijetic. Ofdm for next-generation optical access networks. *Lightwave Technology, Journal of*, 30(4):384–398, Feb 2012. Cited page [61](#)
 - [124] J. Karaki, E. Giacomidis, D. Grot, T. Guillosoy, C. Gosset, R. Le Bidan, T. Le Gall, Y. Jaouën, and E. Pincemin. Dual-polarization multi-band ofdm versus single-carrier dp-qpsk for 100 gb/s long-haul wdm transmission over legacy infrastructure. *Opt. Express*, 21(14):16982–16991, Jul 2013. Cited pages [61](#), [70](#), and [112](#)
 - [125] Sander L. Jansen, Itsuro Morita, Tim C. Schenk, Dirk van den Borne, and Hideaki Tanaka. Optical ofdm—a candidate for future long-haul optical transmission systems. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference*, page OMU3, 2008. Cited page [61](#)
 - [126] Xingwen Yi, William Shieh, and Yiran Ma. Phase noise effects on high spectral efficiency coherent optical ofdm transmission. *J. Lightwave Technol.*, 26(10):1309–1316, May 2008. Cited pages [63](#), [66](#), and [67](#)
 - [127] M. Speth, S.A Fechtel, G. Fock, and H. Meyr. Optimum receiver design for wireless broad-band systems using ofdm. i. *Communications, IEEE Transactions on*, 47(11):1668–1677, Nov 1999. Cited page [64](#)
 - [128] T.M. Schmidl and D.C. Cox. Robust frequency and timing synchronization for ofdm. *Communications, IEEE Transactions on*, 45(12):1613–1621, Dec 1997. Cited page [64](#)
-

BIBLIOGRAPHY

- [129] H. Minn, V.K. Bhargava, and K.B. Letaief. A robust timing and frequency synchronization for ofdm systems. *Wireless Communications, IEEE Transactions on*, 2(4):822–839, July 2003. Cited page [64](#)
- [130] Kai Shi and E. Serpedin. Coarse frame and carrier synchronization of ofdm systems: a new metric and comparison. *Wireless Communications, IEEE Transactions on*, 3(4):1271–1284, July 2004. Cited page [64](#)
- [131] William Shieh, Xingwen Yi, Yiran Ma, and Qi Yang. Coherent optical ofdm: has its time come? *J. Opt. Netw.*, 7(3):234–255, Mar 2008. Cited page [65](#)
- [132] Sander L. Jansen, Itsuro Morita, Tim C. Schenk, and Hideaki Tanaka. Long-haul transmission of 16x52.5 gbits/s polarization-division- multiplexed ofdm enabled by mimo processing. *J. Opt. Netw.*, 7(2):173–182, Feb 2008. Cited pages [65](#), [66](#), and [207](#)
- [133] Xiang Liu and Fred Buchali. Intra-symbol frequency-domain averaging based channel estimation for coherent optical ofdm. *Opt. Express*, 16(26):21944–21957, Dec 2008. Cited pages [66](#), [97](#), and [196](#)
- [134] Sander L. Jansen, Itsuro Morita, Tim C. W. Schenk, Noriyuki Takeda, and Hideaki Tanaka. Coherent optical 25.8-gb/s ofdm transmission over 4160-km ssmf. *J. Lightwave Technol.*, 26(1):6–15, Jan 2008. Cited pages [67](#) and [70](#)
- [135] G.J. Foschini and M.J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3):311–335, 1998. Cited pages [68](#) and [71](#)
- [136] N. Al-Dhahir. FIR channel-shortening equalizers for mimo isi channels. *Communications, IEEE Transactions on*, 49(2):213–218, Feb 2001. Cited pages [70](#) and [152](#)
- [137] Yan Tang, William Shieh, and Brian S. Krongold. Dft-spread ofdm for fiber nonlinearity mitigation. *Photonics Technology Letters, IEEE*, 22(16):1250–1252, Aug 2010. Cited page [70](#)
- [138] Emre Telatar. Capacity of multi-antenna gaussian channels. *European transactions on telecommunications*, 10(6):585–595, 1999. Cited page [71](#)
- [139] Lizhong Zheng and D.N.C. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *Information Theory, IEEE Transactions on*, 49(5):1073–1096, May 2003. Cited page [72](#)
- [140] Vahid Tarokh, N. Seshadri, and AR. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *Information Theory, IEEE Transactions on*, 44(2):744–765, Mar 1998. Cited pages [72](#), [74](#), [75](#), [104](#), [106](#), [192](#), [202](#), and [203](#)
- [141] B. Hassibi and B.M. Hochwald. High-rate codes that are linear in space and time. *Information Theory, IEEE Transactions on*, 48(7):1804–1824, Jul 2002. Cited page [75](#)
- [142] David Tse and Pramod Viswanath. *Fundamentals of wireless communications*. Cambridge university press, 2005. Cited page [75](#)
- [143] J. Boutros and E. Viterbo. Signal space diversity: a power- and bandwidth-efficient diversity technique for the rayleigh fading channel. *Information Theory, IEEE Transactions on*, 44(4):1453–1467, Jul 1998. Cited pages [75](#) and [78](#)
- [144] G. Caire, Giorgio Taricco, and Ezio Biglieri. Bit-interleaved coded modulation. *Information Theory, IEEE Transactions on*, 44(3):927–946, May 1998. Cited page [75](#)
- [145] D. Divsalar and Marvin K. Simon. The design of trellis coded mpsk for fading channels: performance criteria. *Communications, IEEE Transactions on*, 36(9):1004–1012, Sep 1988. Cited page [75](#)

-
- [146] S. Alamouti. A simple transmit diversity technique for wireless communications. *Selected Areas in Communications, IEEE Journal on*, 16(8):1451–1458, Oct 1998. Cited page [77](#)
 - [147] Vahid Tarokh, Hamid Jafarkhani, and AR. Calderbank. Space-time block codes from orthogonal designs. *Information Theory, IEEE Transactions on*, 45(5):1456–1467, Jul 1999. Cited page [78](#)
 - [148] H. El-Gamal and M.-O. Damen. Universal space-time coding. *Information Theory, IEEE Transactions on*, 49(5):1097–1119, May 2003. Cited pages [78](#), [139](#), [140](#), [143](#), [219](#), and [220](#)
 - [149] F. Oggier, G. Rekaya, J. C Belfiore, and E. Viterbo. Perfect space-time block codes. *Information Theory, IEEE Transactions on*, 52(9):3885–3902, Sept 2006. Cited page [79](#)
 - [150] P. Elia, B. A Sethuraman, and P.V. Kumar. Perfect space-time codes for any number of antennas. *Information Theory, IEEE Transactions on*, 53(11):3853–3868, Nov 2007. Cited page [79](#)
 - [151] J. C Belfiore, G. Rekaya, and E. Viterbo. The golden code: a 2 times;2 full-rate space-time code with nonvanishing determinants. *Information Theory, IEEE Transactions on*, 51(4):1432–1436, April 2005. Cited pages [79](#), [192](#), and [198](#)
 - [152] O. Tirkkonen and A Hottinen. Improved mimo performance with non-orthogonal space-time block codes. *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, 2:1122–1126 vol.2, Nov 2001. Cited page [80](#)
 - [153] Ezio Biglieri, Yi Hong, and E. Viterbo. On fast-decodable space-time block codes. *Information Theory, IEEE Transactions on*, 55(2):524–530, Feb 2009. Cited page [80](#)
 - [154] A Mejri, L. Luzzi, and G.R.-B. Othman. On the diversity of the naive lattice decoder. *Systems, Signal Processing and their Applications (WOSSPA), 2011 7th International Workshop on*, pages 379–382, May 2011. Cited pages [83](#) and [221](#)
 - [155] E. Viterbo and J. Boutros. A universal lattice code decoder for fading channels. *Information Theory, IEEE Transactions on*, 45(5):1639–1642, Jul 1999. Cited pages [84](#), [183](#), and [213](#)
 - [156] O. Damen, A. Chkeif, and J. C Belfiore. Lattice code decoder for space-time codes. *Communications Letters, IEEE*, 4(5):161–163, May 2000. Cited pages [84](#), [128](#), [148](#), [183](#), [221](#), and [226](#)
 - [157] E. Agrell, T. Eriksson, A Vardy, and K. Zeger. Closest point search in lattices. *Information Theory, IEEE Transactions on*, 48(8):2201–2214, Aug 2002. Cited page [84](#)
 - [158] AR. Murugan, H. El-Gamal, M.-O. Damen, and G. Caire. A unified framework for tree search decoding: rediscovering the sequential decoder. *Information Theory, IEEE Transactions on*, 52(3):933–953, March 2006. Cited page [84](#)
 - [159] G.R. Ben-Othman, R. Ouertani, and A Salah. The spherical bound stack decoder. *Networking and Communications, 2008. WIMOB '08. IEEE International Conference on Wireless and Mobile Computing*,, pages 322–327, Oct 2008. Cited page [84](#)
 - [160] G.J. Foschini, G.D. Golden, R.A Valenzuela, and P.W. Wolniansky. Simplified processing for high spectral efficiency wireless communication employing multi-element arrays. *Selected Areas in Communications, IEEE Journal on*, 17(11):1841–1852, Nov 1999. Cited pages [86](#) and [225](#)
 - [161] Souha Ben Rayana, Ghaya Rekaya-Ben Othman, Yves Jaouën, and Hichem Besbes. Joint equalization and polarization-time coding detection to mitigate pmd and pdl impairments. *Advanced Photonics Congress*, page SpW2B.3, 2012. Cited pages [88](#) and [89](#)
-

BIBLIOGRAPHY

- [162] Anton Andrusier, Eado Meron, Meir Feder, and Mark Shtaif. Optical implementation of a space-time-trellis code for enhancing the tolerance of systems to polarization-dependent loss. *Opt. Lett.*, 38(2):118–120, Jan 2013. Cited pages [88](#) and [89](#)
- [163] Yan Han and Guifang Li. Coherent optical communication using polarization multiple-input-multiple-output. *Opt. Express*, 13(19):7527–7534, Sep 2005. Cited page [88](#)
- [164] Yan Han and Guifang Li. Polarization diversity transmitter and optical nonlinearity mitigation using polarization-time coding. *Optical Amplifiers and Their Applications/Coherent Optical Technologies and Applications*, page CThC7, 2006. Cited page [88](#)
- [165] Ivan B. Djordjevic, Lei Xu, and Ting Wang. Alamouti-type polarization-time coding incoded-modulation schemes with coherent detection. *Opt. Express*, 16(18):14163–14172, Sep 2008. Cited page [88](#)
- [166] Eado Meron, Alon Andrusier, Meir Feder, and Mark Shtaif. Use of space-time coding in coherent polarization-multiplexed systems suffering from polarization-dependent loss. *Opt. Lett.*, 35(21):3547–3549, Nov 2010. Cited pages [89](#), [103](#), and [192](#)
- [167] Sami Mumtaz, Jingshi Li, Swen Koenig, Yves Jaouen, Rene Schmogrow, Ghaya Rekaya-Ben Othman, and Juerg Leuthold. Experimental demonstration of pdl mitigation using polarization-time coding in pdm-ofdm systems. *Advanced Photonics*, page SPWB6, 2011. Cited pages [102](#), [111](#), [192](#), and [207](#)
- [168] Sami Mumtaz, Ghaya Rekaya-Ben Othman, Yves Jaouen, Jingshi Li, Swen Koenig, Rene Schmogrow, and Juerg Leuthold. Alamouti code against pdl in polarization multiplexed systems. *Advanced Photonics*, page SPTuA2, 2011. Cited pages [89](#), [111](#), and [192](#)
- [169] Jay N Damask. *Polarization optics in telecommunications*, volume 101. Springer, 2005. Cited pages [94](#) and [99](#)
- [170] Ping Lu, Liang Chen, and Xiaoyi Bao. Statistical distribution of polarization-dependent loss in the presence of polarization-mode dispersion in single-mode fibers. *Photonics Technology Letters, IEEE*, 13(5):451–453, May 2001. Cited page [96](#)
- [171] A Iwaki, T. Kawai, Y. Sakamaki, S. Yamamoto, T. Inui, H. Kawakami, S. Yamanaka, T. Ono, K. Mori, M. Suzuki, T. Kataoka, M. Fukutoku, T. Sakano, M. Tomizawa, and Y. Miyamoto. Combined effect of dgd and pdl evaluated in field testbed for 128-gb/s pdm-qpsk signal. *Opto-Electronics and Communications Conference (OECC), 2012 17th*, pages 35–36, July 2012. Cited page [97](#)
- [172] W. Shieh. Pmd-supported coherent optical ofdm systems. *Photonics Technology Letters, IEEE*, 19(3):134–136, Feb 2007. Cited pages [97](#) and [196](#)
- [173] Pierre Delesques, Philippe Ciblat, Gwillerm Froc, Yves Jaouën, and Cédric Ware. Outage probability derivations for pdl-disturbed coherent optical communication. *Advanced Photonics Congress*, page SpTu3A.5, 2012. Cited page [103](#)
- [174] R. Pyndiah, A Glavieux, A Picart, and S. Jacq. Near optimum decoding of product codes. *Global Telecommunications Conference, 1994. GLOBECOM '94. Communications: The Global Bridge., IEEE*, pages 339–343 vol.1, Nov 1994. Cited page [108](#)
- [175] Pierre Delesques, Elie Awwad, Sami Mumtaz, Gwillerm Froc, Philippe Ciblat, Yves Jaouën, Ghaya Rekaya, and Cédric Ware. Mitigation of pdl in coherent optical communications: How close to the fundamental limit? *European Conference and Exhibition on Optical Communication*, page P4.13, 2012. Cited page [111](#)
- [176] Pierre Sillard. Next-generation fibers for space-division-multiplexed transmissions. *40th European Conference and Exposition on Optical Communications*, page Tu.4.1.1, 2014. Cited pages [122](#), [152](#), [160](#), [161](#), [211](#), and [212](#)

- [177] Yasuo Kokubun and Kenichi Iga. Mode analysis of graded-index optical fibers using a scalar wave equation including gradient-index terms and direct numerical integration. *J. Opt. Soc. Am.*, 70(4):388–394, Apr 1980. Cited page [123](#)
- [178] L. Grüner-Nielsen, Y. Sun, R. V. Jensen, J. W. Nicholson, and R.J. Lingle. Splicing of few mode fibers. *Optical Communications, 2014. ECOC 2014. European Conference on*, 2014. Cited pages [132](#) and [133](#)
- [179] S. Alam, Y. Jung, Z. Li, A. Dhar, J.K. Sahu, F. Poletti, and D.J. Richardson. Modal gain equalization in a few moded erbium-doped fiber amplifier. *Photonics Society Summer Topical Meeting Series, 2012 IEEE*, pages 218–219, July 2012. Cited page [133](#)
- [180] F. Ferreira, D. Fonseca, A Lobato, B. Inan, and H. Silva. Reach improvement of mode division multiplexed systems using fiber splices. *Photonics Technology Letters, IEEE*, 25(12):1091–1094, June 2013. Cited pages [134](#) and [215](#)
- [181] Walter F Love. Optical waveguide mode scrambler, October 21 1980. US Patent 4,229,067. Cited page [145](#)
- [182] Masamitsu Tokuda, Shigeyuki Seikai, Koji Yoshida, and Naoya Uchida. Measurement of baseband frequency response of multimode fibre by using a new type of mode scrambler. *Electronics Letters*, 13:146–147(1), March 1977. Cited page [145](#)
- [183] Clemens Koebele, Massimiliano Salsi, Donato Sperti, Patrice Tran, Patrick Brindel, Haik Mardoyan, Sebastien Bigo, Aurélien Boutin, Pierre Sillard, Frederic Cerou, and Gabriel Charlet. Two-mode transmission with digital inter-modal cross-talk mitigation. *37th European Conference and Exposition on Optical Communications*, page Tu.5.B.4, 2011. Cited page [162](#)
- [184] Xi Chen, Jia Ye, Yue Xiao, An Li, Jiayuan He, Qian Hu, and William Shieh. Equalization of two-mode fiber based mimo signals with larger receiver sets. *Opt. Express*, 20(26):B413–B418, Dec 2012. Cited page [162](#)
- [185] D.A. GORE, R.U. Nabar, and A. Paulraj. Selecting an optimal set of transmit antennas for a low rank matrix channel. *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, 5:2785–2788 vol.5, 2000. Cited pages [162](#) and [164](#)
- [186] A.F. Molisch and M.Z. Win. MIMO systems with antenna selection. *Microwave Magazine, IEEE*, 5(1):46–56, Mar 2004. Cited pages [162](#) and [164](#)
- [187] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012. Cited page [177](#)

Publications

Conference proceedings

- “Mode Selection and Larger Set Equalization for Mode-Multiplexed Fiber Transmission Systems”, E.M. Amhoud, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, *IEEE/OSA conference on Optical Fiber Communications (OFC) 2015*, paper Th1D.3, Los Angeles - USA, March 2015.
- “Space-Time Coding and Optimal Scrambling for Mode Multiplexed Optical Fiber Systems”, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, *IEEE International Conference on Communications (ICC) 2015*, London - UK, June 2015.
- “Space-Time Codes for Mode-Multiplexed Optical Fiber Transmission Systems”, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, Y. Frignac, *OSA Advanced Photonics Congress: Signal Processing for Photonic Communications (SPPCom)*, paper SM2D.4, San Diego - USA, July 2014.
- “Polarization-Time coded OFDM for PDL mitigation in long-haul optical transmission systems”, E. Awwad, Y. Jaouën, G. Rekaya-Ben Othman, E. Pincemin, *European Conference and Exhibition on Optical Communication (ECOC)*, P.3.4, London - UK, September 2013.
- “Design criterion of polarization-time codes for optical fiber channels”, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, *IEEE International Conference on Communications (ICC)*, pp.3428-3432, Budapest - Hungary, June 2013.
- “Mitigation of PDL in Coherent Optical Communications: How Close to the Fundamental Limit?”, P. Delesques, E. Awwad, S. Mumtaz, G. Froc, P. Ciblat, Y. Jaouën, G. Rekaya, C. Ware, *European Conference and Exhibition on Optical Communication (ECOC)*, P4.13, Amsterdam - the Netherlands, September 2012.
- “Improving PDL Tolerance of Long-Haul PDM-OFDM Systems Using Polarization-Time Coding ”, E. Awwad, Y. Jaouën, G. Rekaya, *OSA Advanced Photonics Congress: Signal*

Publications

Processing for Photonic Communications (SPPCom), paper SpTu2A.5, Colorado Springs - USA, July 2012.

- “Space-Time Codes for Fiber Communications: Coding Gain and Experimental Validation”, Y. Jaouën, S. Mumtaz, E. Awwad, G. Rekaya-Ben Othman, *IEEE International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP)*, Poznan - Poland, July 2012.
- “100 Gb/s coherent optical polarization multiplexed multi-band-OFDM (MB-OFDM) transmission for long-haul applications”, E. Giacomidis, J. Karaki, E. Pincemin, C. Gosset, R. Le Bidan, E. Awwad, Y. Jaouën, *IEEE International Conference on Transparent Optical Networks (ICTON)*, July 2012.

Journal papers

- “Space-Time Coding Schemes for MDL-Impaired Mode-Multiplexed Fiber Transmission Systems”, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, *submitted to IEEE Journal of Lightwave Technology* in June 2015.
- “Polarization-time coding for PDL mitigation in long-haul PolMux OFDM systems”, E. Awwad, Y. Jaouën, G. Rekaya-Ben Othman, *Optics Express*, vol. 21, pp. 22773-22790, September 2013.

Filed patents

- “Méthode de sélection de modes/coeurs pour transmission sur fibres optiques de type multi-mode/multi-coeur”, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, N° FR 14/58379, Septembre 2014.
- “Méthode et système de transmission sur fibre optique à brassage de modes et/ou de coeurs”, E. Awwad, G. Rekaya-Ben Othman, Y. Jaouën, N° FR 14/56284, Juillet 2014.

Emerging Space-Time Coding Techniques for Optical Fiber Transmission Systems

Elie AWWAD

RÉSUMÉ : La recherche dans le domaine des communications sur fibres optiques avance à un rythme rapide afin de satisfaire des demandes croissantes de communications à débits élevés. Les principaux moteurs de ces avancements sont la multitude de degrés de liberté offerts par la fibre permettant ainsi la transmission de plus de données : l'amplitude, la phase et l'état de polarisation du champ optique, ainsi que le temps et la longueur d'onde sont déjà utilisés dans les systèmes de transmission optique déployés. Pourtant, ces systèmes s'approchent de leur limite fondamentale de capacité et un degré supplémentaire : "la dimension spatiale" est étudié pour réaliser un saut qualitatif majeur en termes de capacité de transmission. Cependant, l'insertion de plusieurs flux de données dans le même canal de propagation induit également des pertes différentielles et de la diaphonie entre les flux, ce qui peut fortement réduire la qualité du système de transmission. Dans cette thèse, nous nous concentrons sur les systèmes de transmission optique de type MIMO basés sur un multiplexage en polarisation ou en modes de propagation. Dans les deux cas, nous évaluons la dégradation de la performance provoquée par une interférence inter-canaux non-unitaire et des disparités de gain entre les canaux engendrées par des imperfections dans les composants optiques utilisés (fibres, amplificateurs, multiplexeurs...), et proposons pour les combattre, de nouvelles techniques de codage pour les systèmes MIMO nommées "codes Spatio-Temporels" (ST), préalablement conçues pour les systèmes radios multi-antennaires.

MOTS-CLEFS : Systèmes à Entrées et Sorties Multiples (MIMO), Communications sur Fibres Optiques, Multiplexage en Polarisation, Multiplexage Spatial, Codage Espace-Temps, Codage Polarisation-Temps, Modulation OFDM.

ABSTRACT : Research in the field of optical fiber communications is advancing at a rapid pace in order to meet the growing needs for higher data rates. The main driving forces behind these advancements are the availability of multiple degrees of freedom in the optical fiber allowing for multiplexing more data : amplitude, phase and polarization state of the optical field, along with time and wavelength are already used in the deployed optical transmission systems. Yet, these systems are approaching their theoretical capacity limits and an extra dimension "space" is investigated to achieve the next capacity leap. However, packing several data channels in the same medium brings with it differential impairments and crosstalk that can seriously deteriorate the performance of the system. In this thesis, we focus on recent optical MIMO schemes based on polarization division multiplexing (PDM) and space division multiplexing (SDM). In both, we assess the performance penalties induced by non-unitary crosstalk and loss disparities among the channels arising from imperfections in the used optical components (fibers, amplifiers, multiplexers...), and suggest novel MIMO coding techniques known as Space-Time (ST) codes, initially designed for wireless multi-antenna channels, to mitigate them.

KEY-WORDS : Multiple-Input-Multiple-Output (MIMO), Optical Fiber Communications, Polarization Division Multiplexing, Spatial Division Multiplexing, Polarization-Time Coding, Space-Time Coding, Orthogonal Frequency Division Multiplexing (OFDM).

