UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Design, monitoring and performance evaluation of high capacity optical networks

Behnam Shariati

Universitat Politècnica de Catalunya

Optical Communications Group

Design, Monitoring, and Performance Evaluation of High Capacity Optical Networks

Behnam Shariati

shariati@ac.upc.edu

This thesis is presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Advisor: Dr. Luis Velasco

Co-advisor: Dr. Jaume Comellas

December 2018

Optical Communications Group (GCO)

Universitat Politècnica de Catalunya (UPC)

C/ Jordi Girona, 1-3

Campus Nord, D4-213

08034 Barcelona, Spain

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola de Doctorat

Acta de calificación de tesis doctoral

Curso académico:

Nombre y apellidos

Programa de doctorado

Unidad estructural responsable del programa

Resolución del Tribunal

Reunido el Tribunal designado a tal efecto, el doctorando / la doctoranda expone el tema de la su tesis doctoral titulada _____

_____.

Acabada la lectura y después de dar respuesta a las cuestiones formuladas por los miembros titulares del tribunal, éste otorga la calificación:

☐ NO APTO ☐ APROBADO ☐ NOTABLE ☐ SOBRESALIENTE

(Nombre, apellidos y firma) Presidente/a	(Nombre, apellidos y firma) Secretario/a	
(Nombre, apellidos y firma) Vocal	(Nombre, apellidos y firma) Vocal	(Nombre, apellidos y firma) Vocal

_____, _____ de _____ de _____

El resultado del escrutinio de los votos emitidos por los miembros titulares del tribunal, efectuado por la Escuela de Doctorado, a instancia de la Comisión de Doctorado de la UPC, otorga la MENCIÓN CUM LAUDE:

☐ SÍ ☐ NO

(Nombre, apellidos y firma) Presidente de la Comisión Permanente de la Escuela de Doctorado	(Nombre, apellidos y firma) Secretaria de la Comisión Permanente de la Escuela de Doctorado

Barcelona a _____ de _____ de _____

Acknowledgment

I could not finish this PhD thesis without invaluable help and support from so many people with whom I have been working or collaborating in the last few years. First of all, I would like to thank my supervisors, Dr. Luis Velasco and Dr. Jaume Comellas, for their guidance, patience, and understanding throughout the course of this PhD. I would also like to thank Dr. Ioannis Tomkos, who acted as my co-supervisor, for all his advice and support during the period I was working in his research group in Athens; it was a pleasant and unique opportunity, which provided me with a great bunch of tools and skills to pursue my career.

In addition, I would like to thank my colleagues, Dr. Marc Ruiz, Dr. Dimitris Klonidis, Dr. José-Manuel Rivas Moscoso, Dr. Pouria Sayyad Khodashenas, and Dr. Nikolaos-Panteleimon Diamantopoulos, for their help and effort to put me back on track whenever I was going in a wrong direction.

I would also like to thank Dr. José-Luis Izquierdo Zaragoza, Dr. Lluís Gifre Renom, Dr. Eduardo Grampín, and Dr. Albert Rafel Porti, who have reviewed this PhD thesis.

Last, but not least, I would really like to thank my family and friends for always being there for me. A special thanks to my dear friend, Pooyan, who has shared quite a lot of time with me to relieve the stress and the tension of this long journey.

This PhD thesis is dedicated to my dear parents, Shamsi and Bagher,
and my dear brothers and sisters; Adel, Fazel, Farah, Reza, and Shahla,
who have taught me how to live and how to love.

Abstract

In the last decade, Internet traffic has seen an enormous growth and it is expected to keep increasing exponentially due to the emergence of vast number of innovative online services and applications. In this regard, optical networks, which are the cornerstone of the underlying Internet infrastructure, have been continuously evolving to carry the ever-increasing traffic in a more *flexible*, cost-*effective*, and *intelligent* way. Having these three targets in mind, throughout the course of this PhD study, I have focused on two general areas for the performance improvement and the evolution of optical networks: *i)* introducing further cognition to the optical layer, and *ii)* introducing new networking solutions focusing on revolutionizing the optical transport infrastructure.

In the first part, we propose Machine Learning (ML) based solutions for performance improvement of Elastic Optical Networks (EON). In this regard, we propose to monitor some Quality of Transmission (QoT) metrics in the optical layer and make operational decisions with the help of sophisticated data analytics approaches, which eventually can contribute to the cost reduction and lowering the required margin in optical networks.

More specifically, we present novel failure detection and identification solutions in the optical layer utilizing the optical spectrum traces captured by cost-effective coarse-granular Optical Spectrum Analyzers (OSA). We demonstrate the effectiveness of the developed solutions for detecting and identifying filter-related failures in the context of Spectrum-Switched Optical Networks (SSON), as well as transmitter-related laser failures in Filter-less Optical Networks (FON).

In addition, at the subsystem level we propose an Autonomic Transmission Agent (ATA), which triggers local or remote transceiver reconfiguration by predicting Bit-Error-Rate (BER) degradation by monitoring State-of-Polarization (SOP) data obtained by coherent receivers.

Regardless of the developed solutions to push further the performance of the currently deployed optical networks beyond their current status, it is expected that the spectral efficiency of the current standard Single-Mode Fiber (SMF) based

optical network approaches the Shannon capacity limits in the near future, and therefore, a completely new paradigm is required to keep with the pace of the current huge traffic increase. In this regard, Space Division Multiplexing (SDM) is proposed as the ultimate solution to address the looming capacity crunch with a reduced cost-per-bit delivered to the end-users. In consequence, I devote the second part of this thesis to investigate different flavors of SDM based optical networks with the aim of finding the best compromise for the realization of a spectrally and spatially flexible optical network.

SDM-based optical networks can be deployed over various types of transmission media. Additionally, due to the extra dimension (i.e., space) introduced in SDM networks, optical switching nodes can support wavelength granularity, space granularity, or a combination of both. In this thesis, I investigate the efficiency of different switching types for the realization of SDM-based optical networks.

In particular, we evaluate the impact of various spectral and spatial switching granularities on the performance of SDM-based optical networks serving different profiles of traffic with the aim of understanding the impact of switching constraints on the overall network performance. In this regard, we consider two different generations of wavelength selective switches (WSS) to reflect the technology limitations on the performance of SDM networks. In addition, we present different designs of colorless direction-less, and Colorless Directionless Contention-less (CDC) Reconfigurable Optical Add/Drop Multiplexers (ROADM) realizing SDM switching schemes and compare their performance in terms of complexity and implementation cost.

Furthermore, with the aim of realizing the benefits and drawbacks of SDM networks over different types of transmission media, we preset a QoT-aware network planning toolbox and perform comparative performance analysis among SDM network based on various types of transmission media. We also analyze the power consumption of Multiple-Input Multiple-Output (MIMO) Digital Signal Processing (DSP) units of transceivers operating over three different types of transmission media.

The results obtained in the second part of the thesis provide a comprehensive outlook to different realizations of SDM-based optical networks and showcases the benefits and drawbacks of different realizations of SDM networks exploited for different use-cases.

Resumen

Se espera que el tráfico de Internet siga aumentando exponencialmente debido a la continua aparición de gran cantidad de aplicaciones innovadoras. Las redes ópticas, que son la piedra angular de la infraestructura de Internet, han evolucionado continuamente para transportar el tráfico cada vez mayor de una manera más flexible, rentable e inteligente. Teniendo en cuenta estos tres objetivos, esta tesis doctoral se centra en dos áreas cruciales para la mejora del rendimiento y la evolución de las redes ópticas: i) introducción de funcionalidades cognitivas en la capa óptica, y ii) introducción de nuevas estructuras de red que revolucionarán el transporte óptico.

En la primera parte, se presentan soluciones novedosas de detección e identificación de fallos en la capa óptica que utilizan trazas de espectro óptico obtenidas mediante analizadores de espectros ópticos (OSA) de baja resolución (y por tanto de coste reducido). Se demuestra la efectividad de las soluciones desarrolladas para detectar e identificar fallos derivados del filtrado imperfecto en las redes ópticas de conmutación de espectro (SSON), así como fallos relacionados con el láser transmisor en redes ópticas sin filtro (FON).

Además, a nivel de subsistema, se propone un Agente de Transmisión Autónomo (ATA), que activa la reconfiguración del transceptor local o remoto al predecir la degradación de la Tasa de Error por Bits (BER), monitorizando el Estado de Polarización (SOP) de la señal recibida en un receptor coherente. Se han desarrollado soluciones para incrementar el rendimiento de las redes ópticas mediante la reducción de los márgenes y la introducción de inteligencia en la administración de los recursos de la red.

Sin embargo, se espera que la eficiencia espectral de las redes ópticas basadas en fibras monomodo (SMF) se acerque al límite de capacidad de Shannon en un futuro próximo, y por tanto, se requiere un nuevo paradigma que permita mantener el crecimiento necesario para soportar el futuro aumento del tráfico. En este sentido, se propone el Multiplexado por División Espacial (SDM) como la solución que permita la continua reducción del coste por bit transmitido ante ése esperado crecimiento del tráfico. En la segunda parte de esta tesis se investigan diferentes

tipos de redes ópticas basadas en SDM con el objetivo de encontrar soluciones para la realización de redes ópticas espectral y espacialmente flexibles.

Las redes ópticas basadas en SDM se pueden implementar utilizando diversos tipos de medios de transmisión. Además, debido a la dimensión adicional (el espacio) introducida en las redes SDM, los nodos de conmutación óptica pueden conmutar longitudes de onda, fibras o una combinación de ambas. Se evalúa el impacto de la conmutación espectral y espacial en el rendimiento de las redes SDM bajo diferentes perfiles de tráfico ofrecido, con el objetivo de comprender el impacto de las restricciones de conmutación en el rendimiento de la red. En este sentido, se consideran dos generaciones diferentes de conmutadores selectivos de longitud de onda (WSS) para reflejar las limitaciones de la tecnología en el rendimiento de las redes SDM. Además, se presentan diferentes diseños de ROADM, independientes de la longitud de onda, de la dirección, y sin contención (CDC) utilizados para la conmutación SDM, y se compara su rendimiento en términos de complejidad y coste.

Además, con el objetivo de cuantificar los beneficios e inconvenientes de las redes SDM, se ha generado una herramienta de planificación de red que prevé la QoT usando diferentes tipos de fibras. También se analiza el consumo de energía de las unidades DSP de los transceptores MIMO operando en redes SDM con tres tipos diferentes de medios de transmisión.

Los resultados obtenidos en esta segunda parte de la tesis proporcionan una perspectiva integral de las redes SDM y muestran los beneficios e inconvenientes de sus diferentes implementaciones.

Table of Contents

Page

Chapter 1 Introduction 1

1.1 Motivation 1

1.2 Goals of the Thesis 3

1.3 Methodology 5

1.4 Thesis Outline 6

1.5 Contributions and References from the Literature 7

Chapter 2 Background 9

2.1 Optical Networks 9

2.1.1 Data Plane 11

2.1.2 Control Plane 14

2.2 Design and Operation of Optical Networks 15

2.3 Optical Performance Monitoring and Failures 17

2.3.1 Optical Performance Monitoring 17

2.3.2 Failures in Optical Networks 19

2.4 Machine Learning Algorithms 21

2.4.1 Decision Trees 21

2.4.2 Support Vector Machines 22

2.4.3 Naïve Bayes 23

2.4.4 Linear Regression 25

2.5 Future Optical Networks 26

2.5.1 Space Division Multiplexing 27
2.5.2 Enabling Technologies for the Realization of SDM 27
2.5.3 Transmission Media 28
2.6 Physical Impairments in SDM Networks 30
2.6.1 Intra-modal Impairments 31
2.6.2 Inter-modal Impairments 34
2.7 Conclusions 35

Chapter 3 Review of the State-of-the-Art 37
3.1 Detection and Identification of Soft-failures in SSON 37
3.2 Detection and Identification of Soft-failures in FON 38
3.3 Autonomic Transmission Exploiting SOP 40
3.4 SDM Switching Technologies 40
3.5 Resource Allocation in SDM Networks 43
3.6 Conclusions 49

Chapter 4 Feature-based Spectrum Monitoring 51
4.1 Motivation and Objectives 51
4.2 OSA for Soft-Failure Detection and Identification 52
4.2.1 Soft-failure Detection, Identification, and Localization 52
4.2.2 Options for Classification using FeX 55
4.3 Illustrative Results 59
4.3.1 VPI Set-up for Data Collection 59
4.3.2 ML-based Classification Comparison 60
4.3.3 Benefits of Using a Single OSA 62
4.3.4 Benefits of Feature Transformation for Classification 63
4.4 Concluding Remarks 66

Chapter 5 Residual-based Spectrum Monitoring 67
5.1 Residual-based Approach for Optical Spectrum Analysis 67
5.2 Facilitating ML Algorithm Deployment Using Residual Signals 71
5.3 Illustrative Results 73

5.3.1 Comparison of Residual-based and Feature-based Approaches 74
5.3.2 The Efficiency of Residual Adaptation Mechanism 77
5.4 Concluding Remarks 79

Chapter 6 Monitoring of Filter-less Optical Networks... 81
6.1 Motivation of Optical Monitoring in FONs 82
6.2 Signal Identification and Classification 85
6.3 Optical Signal Tracking 86
6.3.1 Feature-based Tracking 86
6.3.2 Residual-based Tracking 87
6.4 Illustrative Results 90
6.4.1 PAM4 Scenario 90
6.4.2 QPSK Scenario 93
6.5 Concluding Remarks 96

Chapter 7 SOP Monitoring 99
7.1 Motivation 99
7.2 Autonomic Transmission 100
7.3 SOP-based Pre-FEC BER Prediction 101
7.4 Illustrative Results 103
7.5 Conclusions 106

Chapter 8 Performance Comparison of Different SDM Switching Schemes 107
8.1 Impact of Traffic Profile on the Performance of SDM Networks 108
8.1.1 Simulation Environment and Assumptions 108
8.1.2 Illustrative Results 109
8.2 Impact of Spatial and Spectral Granularity on the Performance of SDM Networks 112
8.2.1 Motivation 112
8.2.2 Simulation Environment and Assumptions 113
8.2.3 Illustrative Results 114

8.3 Conclusions 120

Chapter 9 SDM Networks Based on SMF Bundles, MCFs, and FMFs 123

9.1 Quality of Transmission Estimator 123
9.2 Analytical Models 124
9.2.1 An Analytical Model for FMFs 124
9.2.2 An Analytical Model for MCFs 128
9.2.3 Simulation Environment 129
9.2.4 Transmission Model 131
9.3 Simulations Results 132
9.4 End-to-End QoT Estimator 134
9.5 Network-wide Performance Evaluation 135
9.6 Concluding Remarks 139

Chapter 10 Node Design and Power Consumption Analysis 141

10.1 Comparison of ROADM Architectures for SDM Networks 141
10.1.1 ROADM Architectures and Cost Analysis 143
10.1.2 Results and Discussions 146
10.2 The Impact of MIMO Processing on the Performance of SDM Networks 148
10.2.1 Power Consumption of MIMO Processing for SDM 149
10.2.2 Network-wide Performance Evaluation 152
10.3 Conclusions 153

Chapter 11 Closing Discussion 155

11.1 Main Contributions 155
11.2 List of Publications 156
11.2.1 Publications in Journals 156
11.2.2 Publications in Conferences 157
11.2.3 Book Chapter 159
11.2.4 Other Works not Included in this PhD Thesis 159

11.3 List of Research Projects 160
11.3.1 European Funded Projects 160
11.3.2 National Funded Projects 160
11.3.3 Pre-doctoral Funding Scholarship 161
11.4 Collaborations 161
11.5 Other Achievements 161
11.6 Topics for Further Research 161

List of Acronyms 165

List of References 169

List of Figures

Page

Fig. 1-1. Diagram detailing the followed methodology 6
Fig. 2-1. Heterogeneous optical network architecture 10
Fig. 2-2. SSON (a) and FON (b) 11
Fig. 2-3. Different planes of an optical network 11
Fig. 2-4. Constellation diagram of different modulation formats 13
Fig. 2-5. Functionality of the WSSs 14
Fig. 2-6. MDA architecture 15
Fig. 2-7. Optical spectrum of a QPSK modulated signal 18
Fig. 2-8. An example of polarization rotation in the form of Stokes parameters 18
Fig. 2-9. Optical spectrum for a normal one (a), one experiencing FS (b), and FT (c). 19
Fig. 2-10. The evolution of -3dB bandwidth of the signal w.r.t cascading level (b) . 20
Fig. 2-11. An example of a DT classifier 22
Fig. 2-12. Binary classification problem and the corresponding optimum hyperplane 23
Fig. 2-13. An example of linear regression 26
Fig. 2-14. Different types of transmission media 29
Fig. 3-1. SDM node architecture, for S=4 and D=3, enabling Ind-Sw 41
Fig. 3-2. SDM node architecture enabling a) FrJ-Sw and b) J-Sw 42
Fig. 3-3. Routing principle in the case of J-Sw scheme 44
Fig. 3-4. Routing principle for FrJ-Sw scheme (G=3) without LC support 45
Fig. 3-5. Routing principle for FrJ-Sw scheme (G=3) with LC support 46

Fig. 3-6. Routing principle for Ind-Sw scheme without LC support 46
Fig. 3-7. Routing principle for Ind-Sw scheme with LC support 47
Fig. 3-8. Three resource allocation policies for SDM networks 48
Fig. 4-1. Architecture for failure identification and localization 53
Fig. 4-2. Relevant points of a QPSK modulated signal 53
Fig. 4-3. The evolution of the features w.r.t magnitude of FS (a, c) and FT (b, d) .. 54
Fig. 4-4. The efficiency of the identified features for classification after 2 WSSs (a, b) and after 4 WSSs (c, d) 55
Fig. 4-5. Feature-based approaches: a) multi-classifier, b) single-classifier. 56
Fig. 4-6. Failure detection and identification workflow 57
Fig. 4-7. VPI setup 60
Fig. 4-8. Correction mask of bw_{-3dB} of the setup shown in Fig. 4-7 60
Fig. 4-9. Accuracy at N1 of DT and SVM for FS 61
Fig. 4-10. Average accuracy over N1-N8 of DT and SVM for FS (a) and FT (b). 62
Fig. 4-11. Robustness of the Ml algorithms while one OSA is considered 63
Fig. 4-12. Prediction accuracy of FS (a) and FT (b) estimators. 64
Fig. 4-13. Estimated FS vs estimated FT as a two-dimensional vector space 64
Fig. 4-14. Comparisons of classifiers with and without additional features 65
Fig. 5-1. Soft-failure detection based on residual analysis 68
Fig. 5-2. Synthetically generated expected signals 68
Fig. 5-3. Residual signal calculation 69
Fig. 5-4. Residual based approach workflow considering a normal signal (a), FS (b), and FT (c). 69
Fig. 5-5. Residual-based classification procedure 70
Fig. 5-6. Out-of-field ML training and in-field model adaptation 71
Fig. 5-7. Residual computation and adaptation mechanism 72
Fig. 5-8. Normalization references for the 4^{th} order Gaussian 73
Fig. 5-9. Un-normalized residual (a), normalized w/o adaptation (b), and normalized with adaptation (c) 74
Fig. 5-10. P^{+}_{LH} and P^{+}_{RH} features for failure detection and identification 75
Fig. 5-11. Accuracy of the different approaches 75
Fig. 5-12. Robustness of the different approaches 76

Fig. 5-13. Average node accuracy w.r.t failure magnitude for a) FS and b) FT..... 77
Fig. 5-14. Accuracy per node w.r.t the sequence of cascaded nodes (c). 78
Fig. 6-1. Example of a FON 82
Fig. 6-2. Signal misalignment and overlap 83
Fig. 6-3. Signal classification example (a) and anomaly detection (b) 84
Fig. 6-4. PAM4 optical spectrum a) w/o LD, b) w/ 600 MHz LD 88
Fig. 6-5. Residuals for different LD magnitudes (a) and the evolution of residual features wr.t. the LD (b). 88
Fig. 6-6. QPSK optical spectrum captured by a 1.2 GHz OSA a) w/o LD, b) w/ 2 GHz LD, and c) residuals for different LD magnitudes..... 89
Fig. 6-7. The evolution of residual features w.r.t the magnitude of LD 90
Fig. 6-8. BER degradation vs LD and/or channel spacing for PAM4 systems 91
Fig. 6-9. The tracking accuracy of the PAM4 optical spectrum 92
Fig. 6-10. The tracking accuracy of the PAM4 signal with a) individual signal analysis, and b) contextual analysis 93
Fig. 6-11. Optical spectrum under normal conditions (a) and when LD causes overlapping (b)..... 94
Fig. 6-12. BER degradation vs LD and/or channel spacing for a QPSK system..... 94
Fig. 6-13. $f_{1(-3dB)}$ computation vs OSA resolution(a) and overlap detection vs OSA resolution (b)..... 95
Fig. 6-14. The tracking accuracy of the QPSK optical spectrum 96
Fig. 7-1. Energy saving with tuning FEC mechanism 100
Fig. 7-2. Autonomic transmission and detail of receiver side 101
Fig. 7-3. SOP-based pre-FEC BER prediction 102
Fig. 7-4. Prediction accuracy 104
Fig. 7-5. BER estimation of two different SOP profiles 105
Fig. 7-6. Threshold violation anticipation vs. w 105
Fig. 7-7. Threshold violation anticipation vs. δ 106
Fig. 8-1. CDFs of the assumed traffic profiles with (a) fixed σ=200 Gb/s and μ=700, 1150, and 1600 Gb/s, (b) fixed μ=1248 Gb/s, and σ=48, 96, 192, 384, 768 Gb/s. 109
Fig. 8-2. BP in terms of average number of live connection per A/D nodes 110

Fig. 8-3. BP in terms of standard deviation when μ=1248 Gb/s and average number of live connections per A/D node is 112. 111
Fig. 8-4. BP w.r.t an average number of live connections per A/D node with μ=350 Gb/s and σ=200 Gb/s for J-Sw assuming three levels of spectral granularities. ... 111
Fig. 8-5. Topology of Telefónica Spanish national network. Nodes with A/D capabilities are illustrated with filled grey circles. 113
Fig. 8-6. Average spectral utilization considering current WSS technology w.r.t load. 116
Fig. 8-7. Average data occupancy considering current WSS technology w.r.t load. 117
Fig. 8-8. Average spectral utilization considering finer resolution WSS w.r.t load. 118
Fig. 8-9. Average data occupancy considering finer resolution WSS w.r.t load. ... 119
Fig. 8-10. Average spectrum utilization per link per fiber for J-Sw considering: (a) the current WSS technology and (b) finer resolution WSS. 120
Fig. 9-1. The maximum reach as a function of (a) fiber span length, (b) the number of co-propagating channels, and (c) channel spacing 133
Fig. 9-2. The average spectrum utilization for (a) the actual sized topology (D) and (b) the doubled sized topology (2D) 136
Fig. 9-3. The number of required transceivers for (a) the actual sized topology (D) and (b) the doubled sized topology (2D). 137
Fig. 9-4. The change in the number of TRxs while considering the large-scale topology compared to the small one for every fiber type independently. The total offered load is 900 Tb/s. 138
Fig. 9-5. The change in the number of TRxs considering MCFs/FMFs compared to SMFs for the large-scale topology. Every point is obtained by subtracting the number of required TRxs for BuSMFs case from the target case. 138
Fig. 10-1. Route and select ROADM architectures for Ind-Sw 142
Fig. 10-2. Route and select ROADM architectures for J-Sw with *S*=2 143
Fig. 10-3. Two A/D module architectures for ROADM colorless directionless operation. 144

Fig. 10-4. Two M×N WSS-based A/D module architectures for ROADM CDC operation. 144
Fig. 10-5. Three MCS-based A/D module architectures for ROADM CDC operation. 145
Fig. 10-6. Relative cot of the components w.r.t the cost of 1×9 WSS..... 147
Fig. 10-7. TRC of different ROADM implementations. 147
Fig. 10-8. MIMO-DSP complexity 150
Fig. 10-9. Predicted (a) power consumption per mode and (b) total consumption per 6×100Gb/s module, for 15 nm CMOS. 151
Fig. 10-10. Blocking probability in terms of total offered load considering (a) 100 W and (b) 200 W per 600 Gb/s modules..... 153

List of Tables

Page

Table 1-1: Thesis goals. 5

Table 3-1: State-of-the-art summary 49

Table 4-1 General classification training algorithm pseudocode 58

Table 4-2: Key characteristics and results for the considered approaches for failure detection and identification 66

Table 5-1: Results Comparison 78

Table 5-2: Key characteristics and results for the considered approaches for failure detection and identification 79

Table 6-1: Signal Classification Algorithm 85

Table 7-1: Training algorithm 103

Table 8-1 Values of selected ChBW with the amount of corresponding spectral contents supported by two WSS technologies 113

Table 8-2 Maximum point-to-point transmission reach vs baud rate, for different modulation formats 115

Table 9-1 Effective area [μm^2] for a 6-mode FMF 131

Table 9-2 Differential mode delay [ps/km] between pairs of modes of a 6-mode FMF 131

Table 9-3 FMF parameters for simulations 132

Table 9-4 MCF parameters for simulations 132

Table 10-1 Number of elements and total cost of the ROADM architectures (Fig. 10-2) 143

Table 10-2 Cost of A/D modules 146

Table 10-3 CMOS power dissipation based on Fig. 2 of [Ki11] 150
Table 10-4 Power-limited reach based on total consumption per 600 Gb/s module 152
Table 10-5 Total power consumption per 600 Gb/s module 152

Chapter 1

Introduction

1.1 Motivation

Service operators are continuously looking for solutions enhancing the reliability of their infrastructure and increasing its capacity cost-effectively. Network performance monitoring plays a key role in guaranteeing high reliability and seamless operation of their infrastructure [Ki04]. Performance monitoring requires continuous access to alternative QoT parameters during the lifetime of a lightpath contributing to the visibility level of an operator about their network. The higher the network visibility through collecting QoT metrics it becomes, the higher the network reliability that can be achieved.

In addition to the conventional ways of performance monitoring [Ki04], the advent of coherent receivers has opened up new opportunities for more sophisticated approaches that significantly enhance the visibility of the network [Ha09]. On the other side, the emergence of Software Defined Networking (SDN) allows centralized access to the collected data to produce meaningful information and command proper reconfiguration of the active connections with the aim of satisfying their required quality of service level [Kr15.1]. Bandwidth-variable transceivers and reconfigurable optical add and drop multiplexers (ROADMs) are two key components enabling real-time reconfigurability of the optical transport layer. Note that, the availability of sophisticated data analytics toolboxes, performing efficient data mining and functional decision making, is the key point allowing the exploitation of the full potential of such a platform. The abovementioned procedures, while completed, make a closed operational loop called Observe, Analyse, Act (OAA) loop, upon which the autonomously reconfigurable optical transport network is built.

One of the key tasks of the OAA cycle is to predict the possible threats degrading the QoT of a lightpath to make proper decision in advance of a service disruption. Component failure is one of the reasons threatening the reliability of the active lightpaths. In [APV17.1], the authors proposed several solutions to monitor the performance of lightpaths at the Transponder (Tp) side to verify their proper operation, as well as to detect BER degradations, thus, anticipating connection disruptions. The authors studied several soft failure causes affecting signal QoT, such as laser drift, filter shift, as well as tight filtering and proposed algorithms to detect and identify the most probable failure. Some of these failures happen in the intermediate nodes where optical switching elements are placed and monitoring the signal solely at the egress (or even ingress) node does not allow their localisation. Therefore, monitoring techniques to analyse and evaluate QoT in-line are required.

The previously mentioned soft failures affect the optical spectrum of the lightpaths in a noticeable way, making it a proper measure to characterize the health of a lightpath. In this regard, a new generation of cost-effective OSA deployable in the intermediate nodes is becoming available, which paves the way to develop optical spectrum based performance monitoring solutions. However, data analytics algorithms capable of interpreting the optical spectrum and providing insights on the quality of a lightpath relying on the characteristics of its spectrum are missing. Note that, the development of such algorithms is essential to provide more flexibility to the service providers in monitoring their network performance. The optical spectrum analysis can be ultimately used by algorithms able to identify and localize failures [Ve17.1]. These algorithms can be deployed in the network controller, as well as in nodes' agents, close to the observation points, to reduce the amount of monitoring data to be conveyed to the control/management plane [APV17.1].

The development of an autonomous optical transport network provides numerous benefits for the service operators including higher flexibility, faster repair time, lower operational margins, and, in consequence, reduced maintenance cost. Such strategies contribute greatly to the serviceability of the network and improve its performance and efficiency; however, when it comes to the theoretical and technological limitations, it becomes evident that they mainly complement the performance of the network, rather than pushing the network to go beyond the practical limits. According to recent reports by Cisco and other responsible bodies, due to the emergence of enormous online and cloud-based services such as social networking, online broadcasting of high-definition videos, as well as the exponential growth of Internet of Things based devices, the fundamental spectral efficiency limits of single-mode fibers (SMFs) is rapidly approaching [Wi14].

To address such limitations, over the last couple of years a large number of significant innovations able to offer, in practice, a capacity increase by a factor of around 10-20 (compared to legacy Wavelength Division Multiplexing (WDM)

systems at 10 Gb/s on a 50 GHz grid) have emerged. Initial efforts targeted innovative modulation/coding techniques, novel switching subsystems and routing algorithms supporting flexible frequency allocations, to increase the spectral density/ utilisation in the optical network. This eventually led to the definition of spectrally flexible/elastic optical networks utilising optical Super-Channels (Sp-Ch) together with spectrally flexible/elastic multiplexing schemes and advanced modulation formats, thus enabling the dynamic and adaptive allocation of end-to-end demands with variable connection characteristics (e.g. requested data rates). However, while the spectrally flexible/elastic optical networking approaches can optimise network resources through increased spectral utilization compared to conventional fixed-grid networks, it has limited growth potential due to the nonlinear Shannon limit, which imposes an upper bound on the transport capacity of an SMF within the limited gain bandwidth of C-band Erbium-Doped Fiber Amplifier (EDFA) optical amplifiers. A solution to increase the fiber capacity is to extend the amplification bandwidth. EDFA amplifiers working on the C+L band or Raman-based amplification systems with different designs can increase network capacity by amplifying broader spectral bands compared to conventional C-band EDFA systems.

Even though current trends show that, as a mid-term option, capacity upgrades may occur with the introduction of multi-band systems, the only evident long-term solution to extend the capacity of optical communication systems relies on the use of some form of SDM [Wi14]. The simplest way to achieve spatial multiplexing is to deploy multiple systems in parallel. However, by simply increasing the number of systems, the cost and power consumption also increases linearly [Wi14]. Therefore, component sharing and integration have to be introduced to limit the increase in cost and power consumption. Yet, concrete implementation of optical networks benefitting from the SDM paradigm is still lacking. For optical networks to successfully introduce the SDM capacity multiplier, new switching, routing, and multiplexing hardware have to be introduced [Ma15.1], in addition to network provisioning and resource allocation algorithms and means to mitigate new transmission impairments, as linear and nonlinear crosstalk between spatial channels.

1.2 Goals of the Thesis

Considering the discussions presented in the previous section, this PhD thesis follows two distinct research directions. However, these two distinct research directions eventually complement each other for the realization of the future advanced high capacity optical networks.

On one hand, it tries to explore data analytics approaches to enhance the autonomous operation of optical networks; primarily focusing on the development of optical spectrum based monitoring solutions, and secondly, focusing on the use of

SOP tracking data. On the other side, it reports a comprehensive and comparative study of different flavours to realize spectrally and spatially flexible optical networks exploiting SDM. The development of SDM-based optical networking utilizing different switching schemes and alternative fibers, such as Few-Mode Fiber (FMF) and Multi-Core Fiber (MCF), is investigated, and their benefits and drawbacks regarding network utilisation, solution complexity, and power consumption are explored. Three specific goals are defined to achieve these targets:

G.1 –Optical spectrum monitoring for reliability improvement of optical networks

This goal focuses on the use of optical spectrum monitoring for failure identification and localization in optical networks and consists of two sub-goals:

- **G.1.1**: Optical spectrum monitoring in filtered optical networks based on active photonic switching nodes
- **G.1.2**: Optical spectrum monitoring in filter-less optical networks based on passive photonic switching nodes

G.2 –SOP tracking for adaptive optical transmission

The objective here is to investigate whether applying data analytics to the data gathered from real-time tracking of SOP traces can enhance the optical transmission performance.

G.3 –Investigation of spectrally spatially flexible optical networks

This goal focuses on design and performance evaluations of different flavours of spectrally spatially flexible optical networks and consists of three sub-goals:

- **G.3.1**: Performance evaluation of SDM networks based on bundles of SMFs, MCFs, and FMFs
- **G.3.2**: Alternative SDM switching nodes and their comparative performance evaluations
- **G.3.3**: Power consumption and complexity analysis of SDM nodes

A summary of the goals of the thesis is presented in Table 1-1.

Table 1-1: Thesis goals.

Goals	Sub-goals
G1 **Optical spectrum monitoring for reliability improvement of optical networks**	**G1.1** Optical spectrum monitoring in filtered optical networks based on active photonic switching nodes
	G1.2 Optical spectrum monitoring in filter-less optical networks based on passive photonic switching nodes
G2 **SOP tracking for adaptive optical transmission**	
G3 **Investigation of spectrally spatially flexible optical networks**	**G3.1** Performance evaluation of SDM networks based bundles of SMFs, MCFs, and FMFs
	G3.2 Alternative SDM switching nodes and their comparative performance evaluations
	G3.3 Power consumption and complexity analysis of SDM nodes

1.3 **Methodology**

In this section, the methodology that we follow in this PhD thesis is presented. The beginning of this PhD thesis has started by reviewing the literature and the already published bodies of work to find the areas that have not been yet explored. This approach is followed while a new idea is going to be developed. The already identified goals have been explored in this way. When the idea is determined, the next step is to formulate it as a particular problem to be solved. When a problem is stated, the next step is to solve the problem. In order to address every single problem, a particular approach should be followed, which we generally define it as a solution. A solution should be developed and broke up into smaller ones before proceeding with the required implementations.

The goals of this thesis require a huge amount of algorithm development and implementation in proper platforms. Depending on the nature of the problem, different programming languages and software are used. Considering the first two goals, the data analytics approaches were developed in MATLAB, R or Python, while the transmission testbeds to collect physical layer information, like the

optical spectrum of lightpaths or QoT metrics, such as Optical Signal-to-Noise Ratio (OSNR) and BER, was developed on VPIPhotonics. In order to address the problems defined in goal **G.3**, an offline optical planning tool has been developed in MATLAB. Using all these software and programming languages, the performance of the proposed solutions have been evaluated, and the valuable data produced have been used for dissemination purposes.

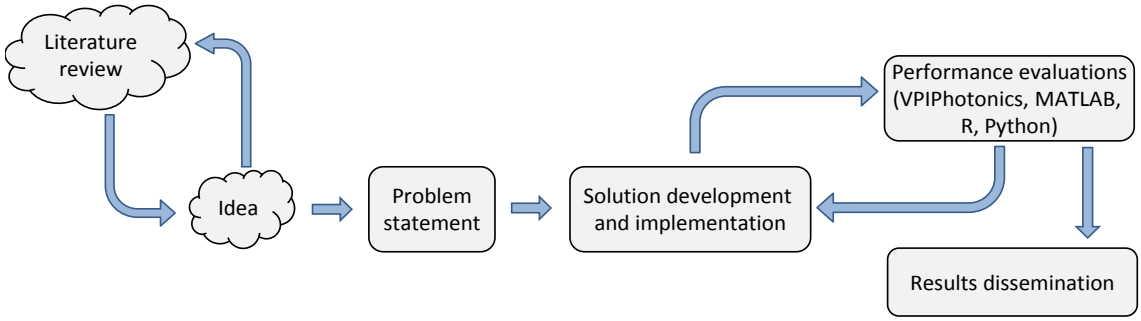


Fig. 1-1. Diagram detailing the followed methodology

1.4 **Thesis Outline**

The remainder of this thesis is organized as follows.

Chapter 2 introduces the background of the different concepts required to follow the studies presented in this PhD thesis.

Chapter 3 reviews the state-of-the-art related to the objectives of this PhD thesis, focusing on identifying and highlighting the niches to be covered.

Chapter 4 presents the studies related to the use of optical spectrum monitoring to detect and identify filter related failures in filtered optical networks. This chapter is based on the following journal and conference publications: [JLT18], [JOCN18.1], [ICTON18.2], [OFC18.2], and [ONDM18].

Chapter 5 continues with the optical spectrum monitoring related studies where an alternative approach is presented to detect and identify filter related failures. This chapter is based on the journal publication [JLT18] and conference publication [OFC19.1].

In Chapter 6, we present spectrum monitoring and signal tracking solutions developed for filter-less networks. This chapter includes the findings presented in two conference papers: [OFC18.1] and [OFC19.2].

In contrast to the previous three chapters, where the main focus is optical spectrum monitoring, in Chapter 7, we present an autonomic transmission agent exploiting the SOP traces collected at the receiver side. This chapter is based on one conference publication [ECOC18].

Moving towards the analysis of future SDM networks, Chapter 8 focuses on a set of comparative network planning analysis to explore the benefits and drawbacks of different SDM switching strategies. The results presented in this chapter summarize the following journal and conference publications: [JLT16], [JLT17], [JOCN17], [ECOC16.1], [ICTON16], and [OFC16.1].

To continue with the SDM related studies, in Chapter 9, we present a set of physical-layer aware performance evaluation studies comparing several SDM network realizations based on different types of fibers. This chapter is mainly based on the materials published in the following journal and conference papers: [JOCN18.2], [ACP16], [ICTON17.1], and [ICTON18.1].

Chapter 10 complements the SDM network studies by presenting a set of complexity and power consumption analysis of some of the solutions presented in the last two chapters. This chapter builds upon the materials presented in the following journal and conference publications: [JLT17], [OFC17.1], [OFC17.2], and [ECOC16.2].

To conclude, Chapter 11 summarizes the main contributions of this PhD thesis, the list of published articles, and the future directions of research.

1.5 Contributions and References from the Literature

For the sake of clarity and readability, references contributing to this PhD thesis are labelled using the following criteria: [<conference/journal acronyms> <Year (*yy*)[.autonum]>], e.g., [OFC17] or [JLT18]; in case of more than one contribution with the same label, a sequence number is added.

The rest of the references to papers or books, both auto references not included in this PhD thesis and other references from the literature are labelled with the initials of the first author's surname and year of publication, e.g., [Ra18]. Finally, references to norms or standards are labeled with its identification, e.g., [RFC7575].

Chapter 2

Background

This chapter presents the main concepts used over the rest of this PhD thesis report.

2.1 Optical Networks

Optical networks are telecommunication networks in, which data is transmitted in the form of light. Most optical networks use optical fibers as the communication channel since they can transmit light with losses as low as 0.2 dB/km and offer broad bandwidth to carry data. In addition to optical fibers, a point-to-point transmission in an optical network requires sophisticated devices to modulate the electrical signal into a lightwave in the transmitter and de-modulate the received optical signal in the receiver side to reproduce the original electrical signal.

In the first generation of optical networks, optical layer was essentially used for transmission and simply to provide capacity, while all the switching, amplification, and other intelligent functions were handled in electronic domain. These networks are known as *opaque* optical networks. An opaque architecture implies that the optical signal undergoes an optical to electronic to optical conversion at different points in the networks; the conversion may happen in the regeneration nodes to re-amplify/re-shape/re-time the signals or in the switching nodes, where the traffic is switched towards its destination.

With the introduction of optical switching nodes and amplification systems, the optical signal could remain in the optical domain from the time it is generated until it leaves the network; these networks are known as *transparent* optical networks. The current networks are mainly deployed in a transparent manner, where optical amplifiers and switches are deployed to manipulate the lightwaves in the optical domain.

In the rest of the thesis, we focus on transparent optical networks. Geographically speaking, optical networks can be categorized in three classes (see Fig. 2-1), mainly depending on their coverage area: *i) access*, which is the closest segment to the end-users, covers different regions in the cities and its suburbs (possibly its longest links span up to 100 km), and usually has more than hundreds of clients to serve, *ii) metro*, which connects different cities (possibly its longest links span up to 200 km), and transport aggregated data of access networks, and *iii) backbone*-or *core*, which connects metro networks in the same country or even national networks.

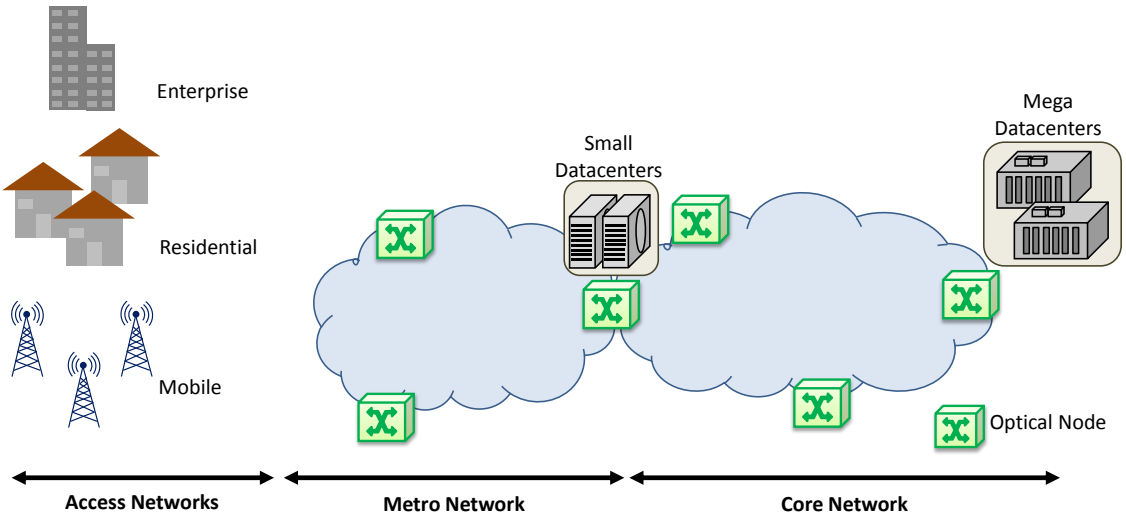


Fig. 2-1. Heterogeneous optical network architecture

The above-mentioned networks mainly benefit from optical switching functionalities in the intermediate nodes, where different lightpaths are routed towards their destinations. To perform per-lightpath routing and switching, optical filtering should be carried out in the intermediate nodes. In this PhD thesis, we call such networks *filtered* or Spectrum-Switched Optical Networks (SSON) shown in Fig. 2-2a. Optical nodes in SSON comprise of Optical Cross Connect (OXC) and Spectrum Selective Switches (SSS). In contrast to the filtered optical networks, we can have Filter-less Optical Networks (FON) in which the nodes are realized just by a pair of splitters and combiners avoiding costly photonic switching devices as shown in Fig. 2-2b. FONs are mainly considered to extend metro and access networks in a cost-effective way. FON nodes comprise of only passive optical splitters and combiners and therefore, they are exploited to create simple network topologies like buses or horseshoes. However, FONs can be combined with OXCs and participate in extending mesh networks based on SSON.

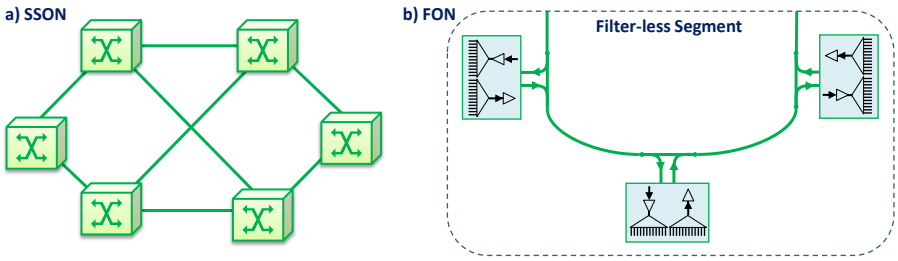


Fig. 2-2. SSON (a) and FON (b)

Generally speaking, optical networks comprise of two different planes: *i) data plane* and *ii) control plane*. The data plane is responsible for transporting the information between different end-nodes, while the control plane is responsible for the coordination of such transportation and related network operations (see Fig. 2-3).

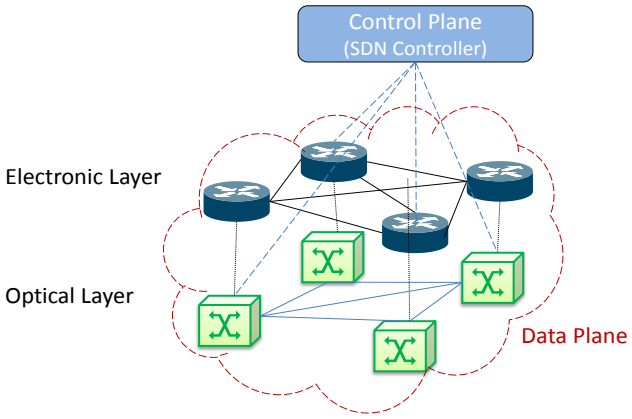


Fig. 2-3. Different planes of an optical network

2.1.1 Data Plane

The data plane includes all the physical components involved in the transmission and switching operation of optical networks. In the following sections, we review the main components and describe their characteristics that allow the proper operation of optical networks.

2.1.1.1 Optical Fibers

Optical fibers are key components for the realization of optical networks. In their simplest form, an optical fiber is composed of a cylindrical core of silica glass covered by a cladding with refractive index (n_2) lower than the refractive index of the core (n_1), which allows the *total internal reflection* of light. We can characterize an optical fiber by two parameters, the relative core-cladding index difference:

$$\Delta = \frac{n_1 - n_2}{n_1} \tag{2-1}$$

and the V parameter, defined as:

$$V = k_0 c_r (n_1^2 - n_2^2)^{1/2} \tag{2-2}$$

where $k_0=2\pi/\lambda$, c_r is the core radius, and λ is the wavelength of the signal [Ag01].

The number of spatial modes supported by a fiber can be determined by the V parameter. Note that, the main difference between a single-mode and a multi-mode fiber is their core size. Optical fibers with V<2.405 supports one single fundamental mode and are called Single-Mode Fiber (SMF). Optical fibers can be designed to support multiple cores (i.e. Multi-Core Fiber (MCF)) or multiple modes (i.e. Multi-Mode Fibers (MMF)). These last types of fibers are further described in Section 2.5.3.

Operationally speaking, an important fiber characteristic is the power loss of a signal during transmission, which is called attenuation constant α and is expressed in dB/km. The attenuation constant depends on the wavelength of the signal. The current low-loss optical fibers have a minimum loss of about 0.2dB/km near 1550 nm.

In addition to the power loss, optical signals traversing over optical fibers suffer from other linear and nonlinear physical impairments that affect the Quality of Transmission (QoT) of the signals. These impairments will be discussed in more detail in Section 2.6.

2.1.1.2 Transceivers

Transceivers (TRx) are used to transmit and receive optical signals over a strand of optical fiber. Transceivers are usually composed of two back-to-back connected transmitter and receiver. An optical transmitter is used to convert an electrical input signal into an optical signal to launch it into the optical fiber, while a receiver is used to detect an optical signal coming from the optical fiber and convert it to a corresponding electrical signal.

The main component of optical transmitters is the optical source, e.g. a laser, while the main component of a receiver is the photodetector. A transceiver is usually characterised by its data rate, which depends on its electrical baud rate and modulation format, and the maximum optical reach, defined as the maximum achievable distance a lightpath can travel without regeneration. Depending on the technology, optical transceivers can generate single-carrier or multi-carrier optical transport channels. While single-carrier optical channels are mainly used in fixed-grid WDM networks, multi-carrier optical transport channels (i.e. super-channels) have been developed to realize flex-grid Elastic Optical Networks (EON) [LoVe16].

There are many modulation formats used in optical transmission systems. Amplitude and phase modulation are the most common techniques used, in which the amplitude or the phase of the optical signal undergo modulation. In this PhD thesis, we work with Binary Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), 8-level Quadrature Amplitude Modulation (8QAM), 16-level Quadrature Amplitude Modulation (16-QAM), and 4-level Pulse Amplitude Modulation (PAM4) formats. Fig. 2-4 shows the constellation diagram of these modulation formats. In addition to modulation to increase the spectral efficiency of optical transmission system, two different polarization modes can be multiplexed; the scheme is known as Dual Polarization (DP) multiplexing. Ultimately, we could even consider both modulation and polarization multiplexing to enhance the spectral efficiency of optical networks.

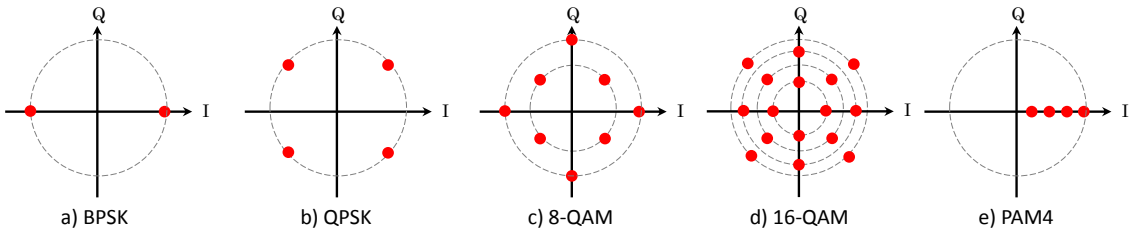


Fig. 2-4. Constellation diagram of different modulation formats

2.1.1.3 Amplifiers

As optical signals experience attenuation while travelling through optical fibers, optical amplifiers are needed to compensate for such attenuation. Even though there are different amplification technologies, Erbium-Doped Fiber Amplifiers (EDFA) are the most common solutions used in the current optical networks, as its optimum performance window coincides with the low-loss region of optical fibers around 1550 *nm*.

Optical amplifiers are not ideal devices and introduce additional noise to the transmission line. Optical amplifiers are characterised by their amplification band, gain profile, amplification flatness in their amplification band, and noise level. Amplified Spontaneous Emission (ASE) noise is the most detrimental source affecting the quality of signals and it is produced due to the imperfection of amplifiers. Considering the characteristics of amplifiers and the underlying fiber infrastructure, their amplification performance can last for particular fiber span lengths (e.g. 60km). Therefore, amplifier span length is considered as a metric determining where to place the amplification nodes in the transmission lines of optical networks.

2.1.1.4 Wavelength Selective Switches

In addition to the components described before, the realization of a transparent optical network requires sophisticated optical switching to be in place. Specifically,

directing wavelength channels through optical networks, e.g. routing at network nodes toward a destination or adding/dropping the channel in the nodes, is performed by Reconfigurable Optical Add/Drop Multiplexers (ROADM). The key element of a ROADM node is the wavelength selective switch (WSS), which performs the tasks of (de)multiplexing and switching on a wavelength basis. WSSs can dynamically route, block, and attenuate all the wavelengths in a fiber. Fig. 2-5 shows an example of a WSS.

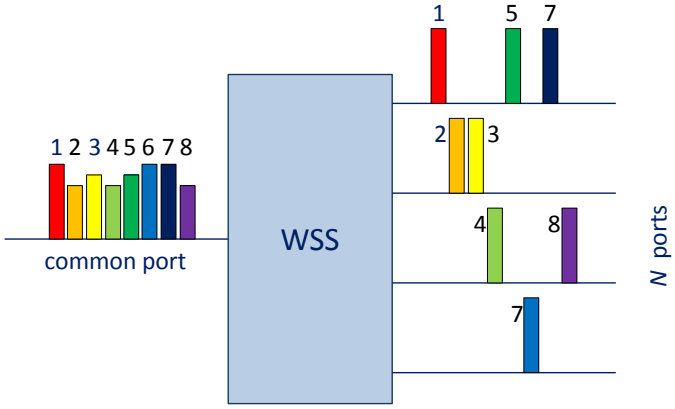


Fig. 2-5. Functionality of the WSSs

WSSs are usually produced in 1×N configuration, meaning that they direct one input port to N output ports as in Fig. 2-5. The wavelength switching process can be dynamically controlled through an electronic communication interface on the WSS. Nowadays, WSSs are mostly based on Liquid-Crystal on Silicon (LCoS); they offer flexible passband filtering with the wavelength switching and routing functionalities [Ma17.1].

WSSs can be characterized by their *spectral switching granularity* and *data occupancy* metric. Current WSS technology allows occupying 32 GHz on a 50 GHz grid; it means they require 18 GHz for guard band. Therefore, the same WSS resolution can allocate finer channels, typically according to a 6.25 GHz grid, i.e., 25.75 GHz can be provisioned on 43.75 GHz, 19.5 GHz on 37.5 GHz, 13.25 GHz on 31.25, or 7 GHz on 25 GHz. The *data occupancy* metric is defined by the bandwidth required to support the data (i.e. equivalent to the baud rate yet measured in GHz) divided by the available bandwidth for data. For example, data occupancy is equivalent to 64%, 52%, and 28% when WSS provision a 32 GHz signal on 50 GHz channel spacing, 19.5 GHz on 37.5 GHz, and 7 on 25 GHz.

2.1.2 Control Plane

While the data plane is responsible for transporting traffic among end-nodes, it is the control plane, which coordinates different components in the data plane to have a network operating properly to deliver the service to the end-users.

In principle, the control plane is in charge of provisioning the available network resources to different services offered by the network; and in an optical network, it is responsible for allocating proper resources to an incoming connection request. It determines the possible route over, which the connection can be provisioned, and subsequently configure the required components to serve that connection. Additionally, it is responsible for reacting against failure in the network by taking proper actions; for instance, re-routing the traffic.

Nowadays, optical networks are under transformation to be controlled by centralized SDN controller. In this new paradigm, the goal is to separate the network's control logic from the underlying routes and switches to promote centralisation of network control and introduce the ability to program the network. Additionally, optical networks can benefit from a separate centralized Monitoring and Data Analytics (MDA) controller which is in charge of managing MDA agents placed in the network nodes (see Fig. 2-6). MDA agents comprise of MDA-enabled modules deployed in the network nodes to perform local data analytics operation [APV17.1].

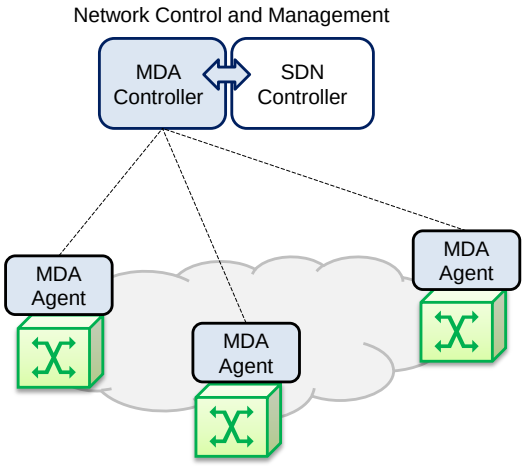


Fig. 2-6. MDA architecture

2.2 Design and Operation of Optical Networks

The design of an optical network requires a set of comprehensive studies to identify the most proper technologies for its realization. These technologies mainly include the fiber infrastructure, transceivers, amplifiers, and optical switching nodes. The identification of the most proper solution for every of these components depend on the characteristics of the components, the target network segment (its geographical size and the expected services to be supported), the behavior of traffic in that segment and its expected growth per year, the considered investment for the deployment, and the expected revenue for the service providers and the operators. Even though Capital Expenditure (Capex) and Operational Expenditure (Opex)

may not be very crucial from an academic point of view, in practice, cost related concerns and the expected amount of revenue are the key decision-making parameters for a solution to win the battle and eventually be deployed in the field.

From a networking perspective, the performance of optical networks is evaluated by carrying out a set of offline network planning practices. In these studies, the objectives are usually defined to minimize the resources to be used for a particular scenario, while keeping the Capex and Opex at a reasonable level to allow service providers and operators to maximise their revenue.

Considering an already defined network topology, i.e. the optical nodes and their interconnected mesh network, and an estimated traffic matrix to be served over such network, offline network planning intend to find the most efficient way to serve all the traffic considering some particular optimization guidelines. These guidelines reflect both technical and economical aspects. Technological aspects include the characteristics and capabilities of the components to be used in the network, as well as their limitations and constraints, while the economical ones include Capex, Opex, and the expected revenue from the services to be provided. The offline network practices include the routing and resource assignment (RSA) problem in the optical network [VeRu17]. The core idea of RSA problems is to find an available wavelength (in a WDM network) or a contiguous portion of the available spectrum (in an EON) on a set of consecutive links, connecting two end-nodes, to establish a lightpath between them.

One of the key targets to consider in the design process is the required QoT that must be guaranteed throughout operation of the network. From the optical layer point of view, there are several parameters to reveal the QoT of the lightpaths; Optical Signal-to-Noise Ratio (OSNR) and Bit-Error-Rate (BER) are the most common metrics used for the representation of QoT. Regardless of the required QoT level, optical networks are designed considering particular margins. Margins are mandatory to ensure that optical networks support the planned demand capacity at during their operation over the full network life [Po17].

At the optical layer, the margin of a lightpath can be quantified as the difference between the actual QoT metric and the threshold above, which the signal is considered recoverable [Po17]. However, the margins come with a cost as they do not allow the full exploitation of the resources of the network. Therefore, there has been an increased interest in decreasing the margins in optical networks and move towards low-margin optical networks [Po17]. In this regard, monitoring the QoT of the active services and lightpaths in the network are required to allow prompt action in case of QoT degradation thus, avoiding full disruption of a connection.

2.3 Optical Performance Monitoring and Failures

In order to guarantee a low-margin optical network designed to operate properly and avoid service level agreement violations, sophisticated monitoring schemes should be employed allowing network operators to take prompt actions in case of possible failures in the network. In addition to the monitoring solutions, appropriate data analytics algorithms should be developed to interpret the collected data facilitating the decision-making process for the operators. In the following, the available monitoring soltuions, which will be later considered in this thesis, are reviewed.

2.3.1 Optical Performance Monitoring

Optical Performance Monitoring (OPM) can enable several important and advanced network functionalities including *i) adaptive impairments compensation, ii) reliable network operation, iii) efficient resource allocation with physical layer consideration*, and *iv) failure detection and identification* [Do15]. Nowadays, with the introduction of coherent transmission systems, OPM has proven to be a key functionality in the receiver side as most of the Digital Signal Processing (DSP) functions built-in the coherent receivers rely on the appropriate estimation of linear impairments, which requires channel parameter monitoring.

OSNR monitoring is commonly realized by the DSP module of the coherent receivers. DSP also monitors pre-Forward Error Correction (FEC) BER. Even though coherent receivers allow monitoring these parameters in the egress node, OSNR and power monitoring is still needed across the network to have full visibility of the network operation and performance. The research conducted in this thesis requires sophisticated approaches for optical spectrum monitoring and State-of-Polarization (SOP) tracking in the networks. These are described in detail in the next sections.

2.3.1.1 Optical Spectrum

Monitoring the optical spectrum of lightpaths requires Optical Spectrum Analyzers (OSA) to be controlled in the networks' nodes. Although, it was not financially justifiable to use such devices as monitoring tools in the optical networks [Finisar.1], with the emergence of compact cost-effective coarse-granular OSA, it is becoming viable to consider this monitoring scheme in the optical networks.

Optical spectrum acquired by OSAs is represented by a set of <frequency, power> pairs, in which the frequency interval of this vector represents the monitoring resolution of OSAs. An example of a 30-GBaud QPSK modulated optical spectrum acquired by an OSA of 312.5 MHz resolution is shown in Fig. 2-7. In general, QPSK and 16QAM optical signals present, once filtered, a flat spectral region around the central frequency, sharp edges, and a round region between the edges and the

central one. Thus, when a signal is properly configured, its central frequency should be around the centre of the assigned frequency slot to avoid filtering effects, and it should be symmetrical concerning its central frequency.

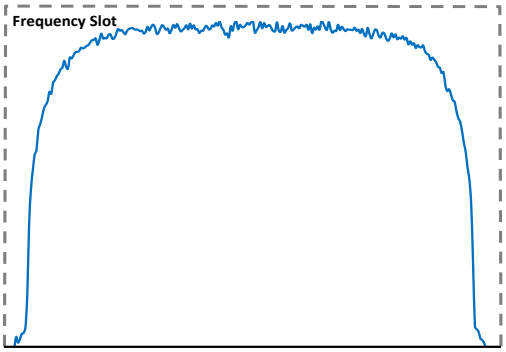


Fig. 2-7. Optical spectrum of a QPSK modulated signal

2.3.1.2 State of Polarization

Optical fibers have specific polarization characteristics. Considering the optical fiber as an optical device, which transmits lights, they can be divided into separated sections that behave like polarization state shifters. Under these circumstances, the evolution of the polarization in the fiber can be described by Stokes vectors and Mueller matrices [Wo98].

SOP represented in the form of Stokes parameters can be monitored to record the rate of polarization rotations [Si17], with the help of DSP units of coherent receivers [Bo18] or a device-based technique [Si17]. SOP traces experiences some level of fluctuations under undesired physical layer effects; so they can be used to anticipate the origin of such undesired phenomenon eventually. Fig. 2-8 shows an example of Stokes parameters representing a random polarization rotation over a period of time.

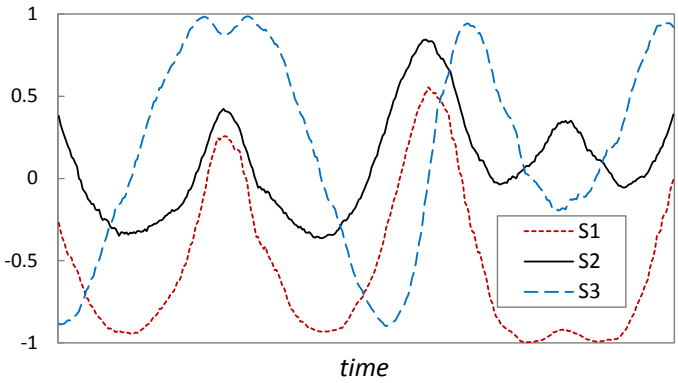


Fig. 2-8. An example of polarization rotation in the form of Stokes parameters

2.3.2 Failures in Optical Networks

In order to have a fully operational optical network, all the components and the underlying infrastructure should operate properly. In this regard, any malfunction, improper operation, or failure that causes an interruption in the network operation must be avoided. Failures in optical networks can be generally categorized in two classes: *i) hard failures*, which are unexpected events that suddenly interrupt established connections [Ma05], like fiber-cut, and *ii) soft-failures*, which are progressive events that degrade the QoT a connection' like Laser Drift (LD), Filter Tightening (FT), and Filter Shift (FS).

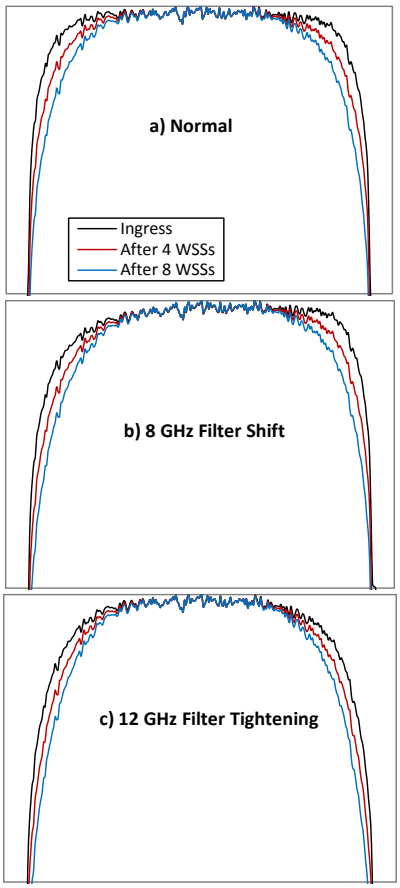


Fig. 2-9. Optical spectrum for a normal one (a), one experiencing FS (b), and FT (c).

As this PhD thesis focuses on the detection and the identification of above-mentioned *soft-failures* with the help of monitoring optical spectrums, we review them in more detail below.

When the signal is properly configured, its central frequency should be around the center of the assigned spectrum slot to avoid filtering effects, and it should be symmetrical with respect to its central frequency. However, any of the following failures results in slight replacement of optical signal or its improper filtering.

- *Laser drift* happens due to the instability in the transmitter, which leads to certain amount of central frequency shift of the lightpath. Such central frequency shift with respect to the assigned slot may result in performance degradation. LD may happen in SSONs or FONs. This phenomenon may result in the *overlapping* of two neighboring signals; a critical problem in FONs. Additionally, in the direct detection systems, LD may result in the *misalignment* of the transmitter and the filter of the receiver.
- *Filter tightening* may arise due to misconfiguration of the filter, which drives the edges of the optical signal to get noticeably rounded. Tight filtering suppresses both ends of a lightpath and results in QoT degradation.
- *Filter shift* makes the optical signal to become asymmetrical with respect to its central frequency. Such asymmetric filtering becomes worse when the signal passes through a cascaded of filters and eventually results in QoT degradation.

Owing to the fact that soft failures might eventually evolve to hard failures, it is of paramount importance not only to detect them a priori before connections disruption but also to localize their cause in order to take proper action, e.g., finding a restoration path for the affected connections avoiding the element in failure [Ve17].

As described in Section 2.1.1.4, WSSs perform optical routing and switching operations at the intermediate nodes in a filtered optical network. When a lightpath passes through a chain of WSSs (which is typical in mesh optical networks where a route comprises of several switching nodes), it experiences a phenomenon known as *filter cascading*. Due to filter cascading, the effective bandwidth (bandwidth at -3dB) of the lightpath reduces. This phenomenon deforms the optical spectra of the lightpaths and has a significant impact on optical spectrum monitoring solutions proposed in this PhD thesis, which will be discussed later in other chapters. For illustrative purposes, the effect of three different levels of filter cascading is shown in Fig. 2-9. Additionally, Fig. 2-10 shows how the effective bandwidth of the signal decreases w.r.t number of cascaded WSSs.

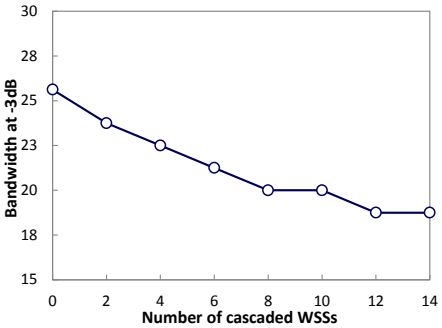


Fig. 2-10. The evolution of -3dB bandwidth of the signal w.r.t cascading level (b)

2.4 Machine Learning Algorithms

Generally, the term Machine Learning (ML) denotes a computer science field grouping algorithms for data analysis able to learn to be used for *classification* tasks, in which a proper category for a new set of observations is identified, and *regression* problems, in which the relationship among given samples is estimated. In other words, in *classification*, the objective is to classify unknown received data, e.g., an optical signal, and decide whether the signal belongs to the normal class, the FS class, or the FT class; while in *regression*, the objective is to predict a behavior; e.g., regression can be used to estimate the magnitude of a failure [Ra18].

ML approaches are usually classified in two distinct learning paradigms: *i) supervised learning*, in which the ML algorithm is first trained with labelled data to learn a general rule that maps inputs to outputs, and *ii) unsupervised learning*, which aims at building representation of a given dataset without receiving labelled data. In this thesis, we take advantage of supervised learning-based ML algorithms; so whenever in the rest of this thesis we refer to ML algorithm, it implies a supervised learning ML algorithm, unless it is otherwise stated.

One of the key steps in the development of a ML solution is feature identification; a set of meaningful features should be identified to characterize the key aspects of the observation under study (e.g. in a simple way, symmetricity and bandwidth could be two meaningful features to characterize an optical spectrum). The next step is to find the most suitable ML algorithm to fit best to our task of interest. Although several ML algorithms might be suitable for the same task, choosing the best one is a problem-dependent decision where their performance needs to be studied for the specific case. Regarding classification, different ML methods are available in the literature, e.g., *Decision Trees* (DT), *Support Vector Machines* (SVM), and *Naïve Bayes* [Ra18.1], [La13]. As for prediction, one of the most popular algorithms is *linear Regression*, which uses observations to find the best polynomial fitting for predictions.

2.4.1 Decision Trees

A decision tree is a supervised learning approach that has a hierarchical tree structure and models the relationships between the features and the potential outcomes. DTs use a structure of branching decisions and leaves that represent the different class labels. Let us describe this approach through the example illustrated in Fig. 2-11, which predicts the class of an optical spectrum capture out of three possible classes: FS, FT, and Normal (N). In our example, the DT is built upon two distinct features: symmetricity (*sym*) and bandwidth of the signal at power -3dB (*bw*).

An observation (i.e. an optical spectrum capture) to be considered begins at the *root node* and it then passes through *decision nodes* that require choices to be made

based on the features of the signal. These choices split the data across different branches following possible outcomes of a decision (yes or no in our example). The tree is terminated by *leaf nodes* when final decisions are made [La13].

A great benefit of DTs is that the resulting structure can be illustrated in a human-readable format. Let's consider the example in Fig. 2-11; starting from the root node, it illustrates if the signal is not symmetric it should belong to the FS class. While it is symmetric, it digs into other aspects of the signal; if the bandwidth of the signal is smaller than a threshold x, then it should belong to the FT class. Any other observation, which is symmetric and its bandwidth is larger than the threshold resided in the leaf belonging to the normal class.

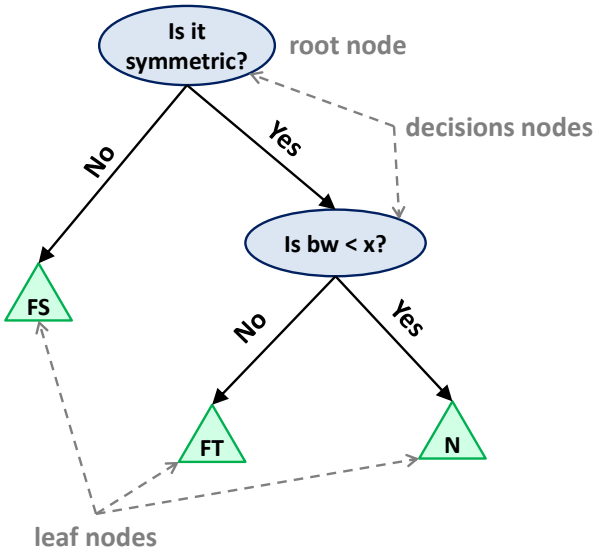


Fig. 2-11. An example of a DT classifier

Typically, DTs are built using *recursive partitioning* algorithm that is a statistical method for multivariable analysis that splits the data into subsets, which are then split repeatedly into smaller subsets until the process stops [La13]. The algorithm usually stops at a node in the case that: *i)* nearly all of the data points at the leaf node have the same class, *ii)* there are no more features to perform splitting, or *iii)* the tree has already reached to its predefined size limit [La13].

There are different parameters to consider while building a DT. The *maximum number of leaves* and *the number of observations per leaf* are the most important ones for the analysis presented in this PhD thesis.

2.4.2 Support Vector Machines

SVM is a supervised learning technique used for binary classification and prediction; in the training phase, the input data is separated into groups of similar features by the computation of a decision boundary, called *hyperplane*. Then, the

SVM considers all the points beyond a particular distance from the hyperplane as one distinct class; that particular distance is called *margin*.

The *margin* is defined as the smallest perpendicular distance between the decision boundary (the hyperplane) and the closest samples [Ra18.1]. Hyperplane is a flat surface in a high-dimensional feature space; however, it becomes a line in a two-dimensional space, as shown in Fig. 2-12. Moreover, as the circles and triangles can be classified perfectly by a straight line, they are called *linearly separable*. However, SVMs can be extended to apply to problems in which the samples are not linearly separable.

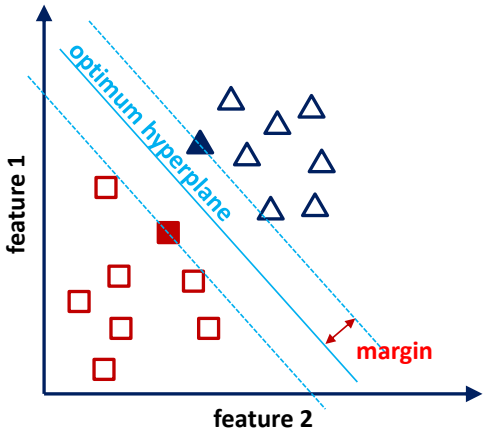


Fig. 2-12. Binary classification problem and the corresponding optimum hyperplane

Considering a training dataset, an SVM can be trained to generate a model for classifying new data samples. Each sample in the dataset belongs to an *n*-dimensional features space. The task of the SVM training is to find a hyperplane that separate the points into two distinct classes. Although there might be several hyperplanes for dividing the points into two distinct classes, the optimum hyperplane, following Maximum Margin Hyperplane (MMH) methodology, is that maximizes the separation distance between the two classes [La13].

When the MMH is found, the samples from each class that are the closest to the MMH are called support vectors (the filled samples in Fig. 2-12). Each class has at least one support vector; however, it is possible to have more than one support vector. The MMH can be defined considering just the support vectors; this is the key potential of SVM approach, which provides a compact way to store a classification model [La13].

2.4.3 Naïve Bayes

Naïve Bayes classifiers are a category of probabilistic classifiers based on Bayes' theorem. These classifiers use the training data to calculate a probability value for each outcome based on the evidences provided by features. Later, when the

classifier is applied to unlabeled data, it considers the probabilities to predict the most likely class.

The Naïve Bayes algorithm is a simple approach in which Bayes' theorem is applied to classification problems. It is called Naïve as it makes some *naïve* assumptions regarding the provided data [La13]; it assumes that all the features in the provided dataset are equally important and independent, which is barely true in realistic problems. However, Naïve Bayes performs well enough even when these assumptions are violated.

To better understand the Naïve Bayes approach, let us first describe the Bayes' theorem; when two events are dependent, they can be described in the following way,

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \tag{2-3}$$

where $P(A|B)$ is the probability of event A given that event B occurred (known as *posterior* probability), $P(B)$ is a measure of how often B is observed to occur in general (evidence probability), while $P(B|A)$ is called *likelihood*, and $P(A)$ the *prior* probability [Bi06].

Considering a dataset that includes observations with q features ($\mathbf{x} = (x_1,\ldots,x_q)$), a Naïve Bayes classification model assigns $P(C_k \mid x_1,\ldots, x_q)$ to each possible classes of C_k. Using Bayes' theorem, we can formulate the model as:

$$P(C_k \mid \boldsymbol{x}) = \frac{P(\boldsymbol{x} \mid C_k)P(C_k)}{P(\boldsymbol{x})} \tag{2-4}$$

In practice, the model solely depends on the numerator as the denominator is independent of the C_k and is constant. Moreover, the likelihood term can be rewritten as,

$$P(\boldsymbol{x} \mid C_k) = P(x_1 \mid x_2,\ldots, x_q, C_k)P(x_2 \mid x_3,\ldots, x_q, C_k)P(x_{q-1} \mid x_q, C_k)P(x_q \mid C_k) \tag{2-5}$$

Now, considering the Naïve assumption that considers every feature to be conditionally independent from other ones, equation 2-5 can be rewritten as:

$$P(\boldsymbol{x} \mid C_k) = \prod_{i=1}^{q} P(x_i \mid C_k) \tag{2-6}$$

Therefore, under the independence assumption, the posterior probability becomes:

$$P(C_k \mid \boldsymbol{x}) = \frac{1}{P(\boldsymbol{x})} P(C_k) \prod_{i=1}^{q} P(x_i \mid C_k) \tag{2-7}$$

where the evidence $P(\boldsymbol{x})$ is a constant and known if the features values are known. Equation 2-7 is the independent feature model, which is the Naïve Bayes probability model. Finally, to have a classifier, we need to combine this model with a rule. A common rule is to pick a hypothesis that is most probable, which is known as Maximum A Posteriori (MAP) decision rule [Bi06].

Considering MAP decision rule, the Naïve Bayes classifier is the function that assigns a class label C_k to some k as follows:

$$\hat{y} = \arg\max P(C_k) \prod_{i=1}^{q} P(x_i \mid C_k), \text{ for } k = \{1, \ldots, K\} \tag{2-8}$$

2.4.4 Linear Regression

Linear regression is a linear method to model the relationship between a dependent variable and one or more independent variables; in this PhD thesis, linear regression is mainly used to estimate an unknown magnitude. In such context, the value to be estimated is the dependent variable, while the features to be used for the estimation are the independent variables.

Mathematically, regression models the data in a slope-intercept form like $y = ax + b$, where y indicates the dependent variable and x indicates the independent variable; the role of linear regression modeling is to find a and b, the slope and the intercept respectively, such that the independent and dependent variables are related with minimum error.

One of the common ways to determine the optimum values of a and b is a method called Ordinary Least Square (OLS) [La13]. In OLS regression, the objective is set to find the slope and the intercept of a line such that the sum of the squared errors is minimized. When the slope and the intercept are found, the model is ready to be used. In case of multiple linear regression modeling with more than one independent variables, the number of slope values to be determined equals to the number of independent variables to be used in the slope-intercept form equation.

Fig. 2-13 shows an example of linear regression. In this example, the goal is to find a model to estimate the magnitude of FS by looking at the symmetricity of the signals. As illustrated, there are several observations representing a linear relation between the symmetricity of the signal and the corresponding value of FS. Therefore, it is possible to find a tuple of <slope, intercept> repressing that relation. When the model is obtained, it can be used for estimating the value of the FS the signal has experienced when its symmetricity is known.

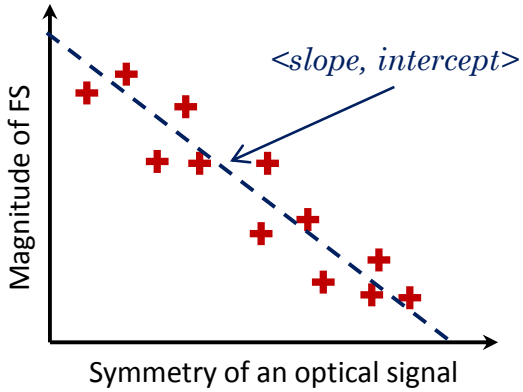


Fig. 2-13. An example of linear regression

2.5 Future Optical Networks

The traffic carried by core optical networks as well as the per-channel interface rates required by IP routers are growing at a remarkable pace year-over-year. Optical transmission and switching advancements have so far satisfied this huge traffic growth by delivering the content over the network infrastructure in a cost and energy efficient manner utilizing to the maximum extent the capabilities of optoelectronic and photonic subsystems and the available bandwidth of deployed optical fibers. However, we are rapidly approaching fundamental spectral efficiency (SE) limits of Shannon channel capacity and the scientific and industrial telecommunications community foresees that the growth capabilities of the current optical networks are quite limited [Wi14].

To address such limitations, over the last couple of years a large number of significant innovations able to offer, in practice, a capacity increase by a factor of around 10-20 (compared to legacy WDM systems at 10 Gb/s on a 50 GHz grid) have emerged. Initial efforts targeted innovative modulation/coding techniques, novel switching subsystems and routing algorithms supporting flexible frequency allocations, to increase the spectral density/utilization in the optical network. This eventually led to the definition of spectrally flexible/elastic optical networks utilizing optical Super-Cannels (Sp-Ch) together with spectrally flexible/elastic multiplexing schemes and advanced modulation formats, thus enabling the dynamic and adaptive allocation of end-to-end demands with variable connection characteristics (e.g. requested data rates). However, while the spectrally flexible/elastic optical networking approaches can optimize network resources through increased spectral utilization compared to conventional fixed-grid networks, it has limited growth potential due to the nonlinear Shannon limit, which imposes an upper bound on the transport capacity of an SMF within the limited gain bandwidth of C-band EDFA optical amplifiers. A solution to increase the fiber capacity is to extend the amplification bandwidth. EDFA amplifiers

working on the C+L band or Raman-based amplification systems with different designs —e.g. all-Raman, hybrid Raman with EDFA amplifiers— can increase network capacity by amplifying broader spectral bands compared to conventional C-band EDFA systems.

Even though current trends show that, as a mid-term option, capacity upgrades may occur with the introduction of multi-band systems, the only evident long-term solution to extend the capacity of optical communication systems relies on the use of some form of Spatial Division Multiplexing (SDM).

2.5.1 Space Division Multiplexing

The simplest way to achieve spatial multiplexing is to deploy multiple systems in parallel. However, by simply increasing the number of systems, the cost and power consumption also increases linearly. To limit the increase in cost and power consumption, component sharing, and integration have to be introduced. To this extent, significant research efforts have focused on the development and performance evaluation of FMFs and MCFs, which can be seen as 'integrated fiber' media, for SDM systems. The development of relevant, flexible optical switches for SDM is an active research field and there are several solutions available today able to perform the switching of spectral/spatial Sp-Chs with variable bandwidth characteristics at a fine granularity, while providing support for all-optical grooming enabling the aggregation and distribution of traffic directly at the optical layer. In addition, significant efforts are being made on the development of the proper control plane framework to orchestrate the operation of such spectrally and spatially flexible networks to bring out their full-potential (i.e. capacity increase and other capabilities such as network virtualization).

2.5.2 Enabling Technologies for the Realization of SDM

A way forward to develop SDM networks is to consider MCF, FMF, Few-Mode MCF (FM-MCF), or even bundles of SMFs, which would allow the network capacity to scale to orders of magnitude higher than what can be achieved with an SMF-based network infrastructure [Wi12], [Wi14], while reducing the cost per bit delivered to the end users by sharing network elements among different spatial dimensions (e.g. integration of multiple optical switching elements [Ma15.1], [Ma17.1], integration of transceivers to form a spatial Sp-Ch [Ri13], or integration of optical amplifiers [Li15] and [Kr12]). Note, however, that the selection of cost-effective technology for each element of an SDM network may result in the introduction of additional physical-layer constraints that need to be taken into account [Ma15.1] and [Ka15].

At early stages, SDM research has mainly focused on SDM fiber technology and transmission performance [Sa16.1], [Mi16], and [Si14.1]. The extra physical impairments introduced by SDM media (compared to the case of SMF

transmission) have been thoroughly investigated [Ka15], e.g., MCF is mostly affected by inter-core Crosstalk (XT), while MMF is strongly impacted by mode coupling and differential mode delay. For this PhD thesis, SDM media can be categorized in three groups, according to whether they have 1) uncoupled/weakly-coupled spatial dimensions (cores, modes or parallel fibers), 2) strongly coupled spatial dimensions, or 3) sub-groups of strongly coupled spatial dimensions [Ma15.1]. These fiber types are further discussed in the next section.

The choice of a transmission medium belonging to one or another of these three categories is a key factor in determining the required architecture and properties of the optical switches at the nodes of an SDM-based network [Ho14.1]. For instance, strongly coupled MCFs or MMFs necessitate that all cores/modes be switched together as the information is mixed, whereas, at the opposite extreme, the use of bundles of SMFs permits both independent and joint switching of the signals on each of the SMFs. A middle-ground solution, e.g. FM-MCF with negligible coupling between cores, would allow independent switching of the mode groups on each of the cores (with a small reach penalty due to the existing, though small, inter-core XT) or joint switching of all modes/cores.

Therefore, switching strategies can be divided into several paradigms, which strongly correlate with the SDM fiber categories defined above [Ma15.1]0[Ma17.1]: *i) Independent Switching* (Ind-Sw), whereby all spectral slices and spatial dimensions can be independently directed to any output port; *ii) Joint Switching* (J-Sw), in which all S spatial dimensions (S is the number of core/modes per SDM fiber) are treated as a single entity, while spectral slices can still be independently switched; and *iii) Fractional Joint Switching* (FrJ-Sw), a hybrid approach in which a number of subgroups of G spatial dimensions out of S (G >1), as well as all spectral slices, can be independently switched to all output ports. The last two paradigms can be categorized as *Spatial Group Switching* (SG-Sw) solutions since the spatial resources are switched in groups rather than independently, as in the case of Ind-Sw. Note that several *spatial switching granularities* result from different levels of grouping of the spatial dimensions: from G=1, which assumes individual fibers, thus corresponding to the Ind-Sw case and offering the finest spatial granularity, all the way through to $G = S$, which considers all spatial dimensions as one spatial group, thus corresponding to the J-Sw case and offering the coarsest spatial granularity. The key enabler for the realization of these switching schemes is the WSS technologies, which will be described in Section 3.3.

2.5.3 Transmission Media

The fundamental concept of SDM relies on placing numerous spatial channels in a single fiber structure aiming at manufacturing an integrated fiber carrying manifold channels in a denser, lighter, and more cost-effective cable. The type of channels depends on the way that SDM is exploited: multiple SMFs in a fiber

bundle, many cores in a single cladding, multiplexed Linearly-Polarized (LP) modes, or several cores each supporting few multiplexed LP modes [Na15]. Even though many SDM fiber alternatives with different number of cores/modes have been developed for the future SDM networks, for this PhD thesis, SDM media can be categorized in three groups, according to whether they have 1) uncoupled/weakly-coupled spatial dimensions (cores, modes or parallel fibers), 2) strongly coupled spatial dimensions, or 3) sub-groups of strongly coupled spatial dimensions [Ma15.1]. In the next couple of paragraphs, three fiber classes are discussed in more detail.

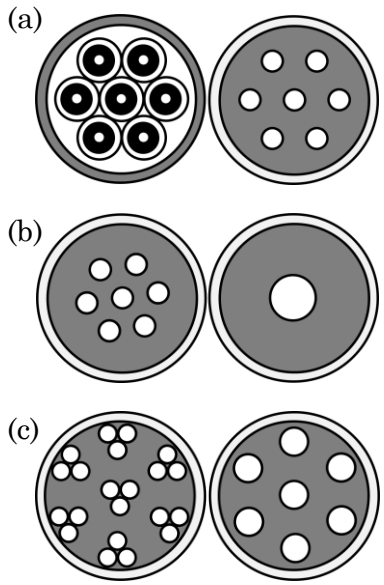


Fig. 2-14. Different types of transmission media

2.5.3.1 Uncoupled/weakly-coupled Spatial Dimensions

Spatial channels of this type of fibers remain distinct in fiber propagation (i.e. there is no/negligible XT between spatial channels) as would be experienced in bundles of SMFs (BuSMFs) or weakly-coupled MCF. Therefore, individual spatial channels can be freely switched from one link to another, or Add/Drop (A/D) operations can be performed independently for any wavelength on any spatial channel. When these fibers are in place, it is also possible to route a wavelength channel from one spatial channel to another —the so-called Lane Changes (LC) method of functioning of switching infrastructure [Ma15.1]—, eliminating wavelength continuity constraint. These types of fibers are designed to avoid any complex Multi-Input Multi-Output (MIMO) processing at the receiver side. However, they require more complex and costly switching infrastructure as we will see in the next section. Fig. 2-14 (a) shows a simple illustration of a cross-section of a bundle of 7 SMFs and a 7-core weakly-coupled MCF.

2.5.3.2 Strongly-coupled Spatial Dimensions

Spatial channels of this type of fibers strongly mix in fiber transmission, as occurs in FMFs and coupled MCFs. Coupled MCFs and FMFs are affected by inter-modal impairments, e.g. MCFs are affected by inter-core XT and FMFs are impacted by mode coupling, differential group delay, mode dependent loss and inter-modal nonlinearity. As a consequence, MIMO processing is required to compensate for these impairments [Ry12]. Since the information is mixed across all spatial modes/cores, spatial channels must be switched together to other destinations, and LC operation is not allowed. Even though this constraint sacrifices the routing flexibility, it significantly simplifies the switching infrastructure and results in huge cost savings. Fig. 2-14(b) shows a cross-section of a 7-core coupled MCF, a single-core FMF carrying several LP modes.

2.5.3.3 Sub-groups of Strongly-coupled Spatial Dimensions

Spatial channels of this type of fibers mix only within the subgroups, as would be experienced in FM-MCF or SDM fibers with uncoupled subgroups of coupled cores. Subgroups are defined by the fiber design, e.g. uncoupled cores of an FM-MCF, and the spatial channels are belonging to a subgroup must be switched as a single entity. Subdividing the spatial channels into smaller size eases the switching limitations concerning the former type of SDM fibers, while not reaching the full flexibility of the first type (i.e. fibers with uncoupled spatial dimensions). Even though routing a wavelength channel from one spatial channel to another within a group is not allowed, routing a wavelength channel from one spatial group to another is applicable assuming that appropriate switching infrastructure is in place. Note that, wherever LC is discussed in the rest of this chapter for this group of fibers, it indicates spatial group changes. Fig. 2-14(c) shows a cross-section of an FM-MCF and an SDM fiber having uncoupled subgroups of coupled cores.

2.6 Physical Impairments in SDM Networks

The physical impairments introduced by SDM fibers are classified in two groups: i) intra-modal impairments and ii) inter-modal impairments. All transmission media are affected by intra-modal linear and nonlinear impairments. Furthermore, MCFs and FMFs also present inter-modal impairments, e.g. MCFs are affected by inter-core XT and FMFs are impacted by mode coupling, Differential Group Delay (DGD), Mode Dependent Losses (MDL), and inter-modal nonlinearities. DSP techniques, including MIMO processing, have been developed in order to compensate for the abovementioned inter-modal linear impairments [Ar14], though they may result in a significant increase in power consumption of the transceiver. Nonlinear impairments, nonetheless, are unlikely to be compensated for with the current DSP modules and, therefore, could result in significant performance

degradation. We will review the most important impairments, which affect the QoT of lightpaths travelling over SDM fibers next.

2.6.1 Intra-modal Impairments

We first start by reviewing the intra-modal impairments. Intra-modal impairments refer to the entire phenomenon that affects the characteristic of an optical signal while passing through a single mode optical fiber; it can be either the fundamental mode of a SMF or one of the guided modes in FMFs. However, the descriptions are focused on SMFs. Intra-modal impairments can be generally categorized in two groups; i) *linear impairments* (such as power loss, Chromatic Dispersion (CD), and Polarization Mode Dispersion (PMD)) and ii) *nonlinear impairments* (such as Self-Phase Modulation (SPM), Cross-Phase Modulation (XPM), and Four-Wave Mixing (FWM)). Power loss, CD, and PMD are linear impairments and can be very well compensated with the help of optical amplifiers and coherent receivers [Ra10]. However, nonlinear impairments are unlikely to be compensated and will result in performance degradation of optical signals.

There are two categories of nonlinear effects. The first arises due to the interaction of light waves with phonons in the silica [Ag01], while the second set of nonlinear effects arises due to the dependence of the refractive index on the intensity of the applied electric field, which itself is proportional to the square of the field amplitude.

In this PhD thesis, we are interested in the second category, which are described next. The most important nonlinear effects in this category are SPM, XPM, and FWM. The nonlinear interaction depends on the transmission length and the cross-sectional area of the fiber. Therefore, the nonlinear effects become worse for longer links. However, with the propagation of the signal along the link, its power decreases because of fiber attenuation. Therefore, most of the nonlinear effects occur in the beginnings of the fiber span and becomes negligible as the signal propagates. Below, we will review the most important impairments.

2.6.1.1 Power Loss

An important fiber parameter is a measure of power loss during transmission of optical signals. There are several factors that contribute to the power loss, with material absorption and Rayleigh scattering as dominant contributors [Ra10]. There are other factors, which contribute to power losses, like scattering of light at the core-cladding interface and bending of fiber. Typically used fibers show a loss of ≈ 0.2 dB/km around 1550 nm. However, the total loss of fiber cables used in real networks is slightly larger (~ 0.03 dB/km) because of splice and cabling loss [Ag01].

2.6.1.2 Chromatic Dispersion

CD refers to a phenomenon in which the different spectral components of a pulse travel at different velocities. CD arises for two reasons. The first is that the refractive index of silica, the manufacturing material of the optical fiber, depends on the frequency. Therefore various components of frequency travel at different speed in silica. This reason results in a component of CD called *material dispersion* [Ra10].

Even though material dispersion is the principal component of CD for most fibers, there is a second component, called *waveguide dispersion*. The light energy of a mode usually propagates partly in the core and partly in the cladding. In addition, the distribution of the power of a mode between the cladding and the core of the fiber is itself wavelength dependent. Therefore, even in the absence of material dispersion, if the wavelength changes, this power distribution changes and causes the effective index or propagation constant of the mode to change resulting in CD.

Fiber dispersion has a critical role in propagation of optical pulses (mainly short ones) because different spectral components associated with the pulse travel at different speeds given by $c / n(\omega)$. Mathematically speaking, fiber dispersion can be expressed by expanding the mode-propagation constant β in a Taylor series around the frequency ω_0 at which the pulse spectrum is centered [Ag01]. In addition, the parameters β_1 and β_2 of the Taylor series are related to the refractive index n and its derivative as follows

$$\beta_1 = \frac{1}{v_g} = \frac{n_g}{c} = \frac{1}{c}\left(n + \omega \frac{dn}{d\omega} \right) \tag{2-9}$$

$$\beta_2 = \frac{1}{c}\left(2\frac{dn}{d\omega} + \omega \frac{d^2 n}{d\omega^2} \right) \tag{2-10}$$

Where v_g is the *group velocity* and n_g is the group index.

Physically speaking, the envelope of an optical pulse travels at the group velocity and the parameter β_2 represents dispersion of the group velocity, which is responsible for broadening of the optical pulse. This phenomenon is known as the Group-Velocity Dispersion (GVD). Finally, to obtain the total dispersion, the waveguide contribution must be added to the material contribution. The total dispersion of a SMF, which is commonly shown by dispersion parameter D is related to β_2 by the relation [Ag01]

$$D = \frac{d\beta_1}{d\lambda} = -\frac{2\pi c}{\lambda^2}\beta_2 \approx \frac{\lambda}{c}\frac{d^2 n}{d\lambda^2} \tag{2-11}$$

2.6.1.3 Polarization-Mode Dispersion

A SMF is not truly single mode because it can support two degenerate modes that are polarized in two orthogonal directions [Ra10]. Under ideal conditions, a mode launched with its polarization in the x direction would not couple to the mode with the orthogonal y polarization state [Ag01]. In real fibers, small deviations from cylindrical symmetry of the fiber can result in the mixing of the two polarization states by breaking the mode degeneracy. Mathematically speaking, the mode propagation constant β becomes a bit different for the modes polarized in the x and y directions. This characteristic is referred to as modal birefringence.

In standard optical fiber, the strength of modal birefringence is not constant along the fiber and changes randomly because of fluctuations in the core shape and stress. Therefore, the polarization of light launched into the fiber with a fixed state of polarization changes in a random fashion. This change is an issue for optical communication systems, mainly when short pulses are transmitted over long distances. Now, if an input pulse is launched on both polarization components, the two components travel travels at different speed along the fiber because of their different group velocities. Thus, the pulse becomes broader at the end because group velocities change randomly w.r.t random changes in fiber birefringence. This phenomenon is called PMD. PMD becomes a limiting factor for high-speed communication systems designed to operate over long distances near the dispersion-tailored wavelength of the fiber [Ra10].

2.6.1.4 Self-Phase Modulation

SPM emerges due to the intensity-dependent component of the refractive index of the fiber. As a result, a phase shift that is proportional to the intensity of the pulse is introduced by this nonlinear refractive index and different parts of the pulse experience varying phase shifts, which give rise to chirping of the pulses. This chirping effect is proportional to the transmitted signal power so that SPM effects are stronger in system using high transmitted powers. The pulse broadening effects of CD is affected by SPM-induced chirp affects and is important to be considered for high-bit-rate systems that also suffer from CD limitations [Ra10].

2.6.1.5 Cross-Phase Modulation

XPM arises when two optical fields propagate at the same time and affect each other through the intensity dependence of the refractive index. The nonlinear coupling induced by XPM can occur not only when two beams of different wavelengths are present on the fiber but also through the interaction between the orthogonally polarized components of a single beam in a birefringent fiber [Ra10].

In WDM systems, the intensity-dependent nonlinear effects are magnified as the combined signal from all the channels can be quite strong, even when individual channels are operated at reasonably moderate powers. Thus, the phase shift arises from the intensity-dependency and the resulting chirping, induced by SPM alone is

enhanced because of the power of the signals in the other channels. This effect is referred to as XPM.

2.6.1.6 Four-Wave Mixing

In a WDM system which operate over the angular frequencies $\omega_1,\ldots, \omega_n$, the intensity dependence of the refractive index not only introduces phase shifts within a channel but also generates signals at new frequencies such as $2\omega_i - \omega_j$ and $\omega_i + \omega_j - \omega_k$ [Ra10]. This phenomenon is called FWM. These signals appear as crosstalk to the existing signals in the system; this crosstalk is particularly severe when the channel spacing is tight. In contrast to SPM and XPM effects, which are significant mainly for high-bit-rate systems, the FWM depends on the bit rate but is highly dependent on the channel spacing and fiber CD. Decreasing the channel spacing or CD can increase the FWM effect. Thus the effects of FWM must be considered even for moderate-bit-rate systems when the channels are closely spaced. In a WDM system, if another channel happens to be located at where the FWM-generated signal appears, the FWM process can significantly degrade the quality of that channel. However, in practice, the signals generated by FWM have lower powers and the attenuation of signals due to fiber loss.

2.6.2 Inter-modal Impairments

In addition to the intra-modal impairments described above, SDM fibers suffer from inter-modal impairments. PMD, which was discussed before, is a particular type of inter-modal impairments, where the impairment arises because the two polarization modes gets coupled due to the imperfection of the cylindrical symmetry of the medium and, as a consequence, the two polarized modes travel with different velocities inside the medium. The rest of this section is dedicated to reviewing the inter-modal impairments, which affect the QoT of a signal traveling through MMF and MCFs.

Typical MMFs have a core larger than a wavelength of light and carriers hundreds of *waveguide modes*. SMF has a core of similar size to a wavelength that restricts SMF to support a single "fundamental" spatial mode. A *mode* is defined as a pattern of the optical electric field that propagates inside the medium without change apart from variations in its amplitude and phase [Ag01]. In contrast to MMF, which carriers hundreds of modes, there are another type of fiber, the so-called FMF, which are designed to carry few guided modes up to a few tens, where each mode is treated as an individual data channel to carry data.

2.6.2.1 Intermodal Dispersion

FMF supports propagation of several spatial modes with different group velocities, and thus different Group Delays (GD), a phenomenon called (*inter*)*modal dispersion*. In the other words, the velocity of the pulse envelope, which determines the pulse distortion in a FMF, is determined mainly by the mode dispersion in

FMF and its derivative with respect to the frequency, the so-called mode group delay (MGD). Modal dispersion in FMF typically results in a larger GD spread than that caused by CD. In most situations, dispersion leads to broadening of signal pulses, which correspond to data bits. In a communication system, this leads to the overlap of pulses representing adjacent bits, distorting the signal [Ra10].

2.6.2.2 Mode Dependent Loss

Transmission fiber and inline optical amplifiers can introduce mode-dependent loss or gain, which are collectively refer to as MDL. MDL causes random variations of the powers of signals propagating in different modes [Ho12], [Ho13]. These power variations may affect the various frequency components of each signal in a different way, and may change over time. Similar to multipath fading in a MIMO wireless system, these power variations can cause MIMO system capacity to become a random variable; thus, the mean capacity may be reduced. Strong mode coupling can reduce the power variations and the associated capacity fluctuations.

2.6.2.3 Mode Coupling

Signals traveling over FMFs suffer from mode coupling. Mode coupling occurs when several signals traveling over different modes get mixed due to energy exchange among them [Ho13]. Even if a signal is launched into one spatial mode, bends, refractive index imperfections and other perturbations cause the signal to couple into multiple modes. Mode coupling can be classified as weak or strong, which depends on whether the total system length is comparable to, or much longer than, a length scale over which propagating fields remain correlated [Ho13].

The characteristics of a FMF, in particular the MGD profile and loss profile vary along the length of a fiber, and can be considered constant only over a certain correlation length. When a fiber is much longer than its corresponding correlation length, it can be modeled as a concatenation of several sections with independent characteristics. This is referred to as the *strong-coupling regime*. In contrast in the *weak-coupling* regime, the correlation length is comparable to the total system length. In this regime, signal propagation can be modeled using a small number of sections, where each section should be a little bit larger than the correlation length of the fiber. Mode coupling leads to crosstalk between spatially multiplexed signals and significantly affects the end-to-end GD spread of a system [Ho13].

2.7 Conclusions

This chapter has been devoted to review the required background to facilitate the understanding of the fundamental concepts used in this PhD thesis. We have started by introducing optical networks and the required components to realize it. Then, we have discussed the design process and operation of such networks. We have moved forward to present the optical performance monitoring concept,

covering mainly optical spectrum and SOP monitoring techniques. Later, we have explained the possible soft-failures they may arise in optical networks and provide background on several ML-based techniques that will be further used in the thesis. Finally, we have presented required background on future SDM networks, the enabling technologies, including transmission media and the switching strategies, and ultimately the physical layer impairments affecting the QoT of signal traversing over SDM fibers.

The following chapter focuses on reviewing the state-of-the-art of the objectives of this PhD thesis, aiming at identifying research opportunities.

Chapter 3

Review of the State-of-the-Art

In this chapter, we review the state-of-the-art related to the different goals defined in Chapter 1 for this PhD thesis, with the two fold objectives of ensuring that they have not yet been covered in previous works in the literature and to serve as a starting point for this research work targeting to occupy the discovered niches.

3.1 Detection and Identification of Soft-failures in SSON

Network performance monitoring [Do15] is a key enabler for failure identification and localization, which can greatly bring down both, the repair time and the Opex cost of optical networks [Lo16]. Many research efforts have been dedicated to developing failure localization techniques for *hard failures*, i.e., unexpected events that suddenly interrupt the established connections (see, e.g., [Ma05]). Nonetheless, the identification and localization of *soft failures*, i.e., events that progressively degrade the QoT, remains rather unexplored. Owing to the fact that soft failures might eventually evolve to hard failures, it is of paramount importance not only to detect them a priori, before connections disruption, but also to localize their cause in order to take the proper action, e.g., finding a restoration path for the affected connections avoiding the element in failure [Ve17].

Such performance monitoring is enabled by the ability of optical components to take measurements. One of the key features to be exploited in next-generation optical networks is the availability of monitoring parameters that can be used by data analytics applications, especially those based on ML [Ra18]. This paves the way for future autonomic networking as defined in [RFC7575], including self-protection and self-healing.

Aiming at detecting traffic anomalies in packet networks, the authors in [APV17.1] proposed bringing data analytics toward the network nodes to reduce the amount

of monitoring data to be conveyed to the control and management plane, while improving detection times. Following such idea, the authors in [Ve18.1] and [Gi18] proposed a distributed Monitoring and Data Analytics (MDA) framework that includes MDA agents running close to the observation points in the network nodes, as well as a centralized MDA controller running in the control and management plane besides the SDN controller. Such MDA framework is the base to build autonomic optical networks, especially in the case of utilising white-boxes, which might include specific optical monitoring devices [Ve18.2].

Recently, the authors in [APV17.2] proposed several solutions to monitor the performance of lightpaths at the Tps side to verify their proper operation, as well as to detect BER degradations, thus, anticipating connection disruptions. The authors studied several soft-failure causes affecting signal QoT, such as those defined in Chapter 2 and proposed algorithms to detect and identify the most probable failure. Some of these failures happen in the optical switching intermediate nodes, so monitoring the signal solely at the end nodes does not allow their localisation. Therefore, monitoring techniques to analyze and evaluate QoT in-line are required.

As the abovementioned failures noticeably affect the optical spectrum of the lightpaths, OSAs can be used to monitor the spectrum along the transmission line aiming at detecting and localizing those type of failures. Practically speaking, the realization of such solutions become possible with the emergence of a new generation of compact, cost-effective OSAs with sub-GHz resolution in the form of optical components [Finisar.1] allowing real-time monitoring of the optical spectrum of the lightpaths and their corresponding OSNR.

One of the challenging tasks is to distinguish whether a signal is affected by an FT failure or it is just degraded due to filter cascading effects, as both phenomena show similar effects. Therefore, aiming at developing sufficiently robust solutions, it is essential to cope with this issue thus, preventing the misclassification of a properly configured signal as a failed one.

In conclusion, there is a lot of room for soft-failure detection and identification at the optical layer by monitoring the data from OSAs.

3.2 Detection and Identification of Soft-failures in FON

FON have recently attracted significant attention as a cost-effective metro solution to interconnect 100G coherent-based nodes in a drop and waste network architecture [Za13]. FONs can also perform very well in small size regional and submarine transmission networks. In contrast to filtered optical networks, where the signals are dropped at their destination, they continue to spread over the

transmission line in FON. This fact leads to spectrum waste and efficiency penalties. On the other hand, since the operating lightpaths do not pass through filtering nodes, FONs can be considered as a kind of *gridless* network where frequency slots are not rigidly defined.

The inherent *gridless* potential of FON could allow the channels to be placed very close to each other aiming to alleviate spectrum waste. Nonetheless, the bottleneck is that (un)intentional LD of a transponder, possibly due to failure or its misconfiguration, might lead to overlapping of neighboring channels. This contrasts with Spectrally-Switched Optical Network (SSON), where LD effects are much more moderate. In addition to the overlapping issue, LD could have a detrimental impact on the lightpath itself due to detuning w.r.t. the receiver optical filter. While Coherent Detection (CoD) receivers can easily track the central frequency of the transmitted signal by evaluating the offset with local oscillator, Direct Detection (DD) systems detect only the intensity of the optical signal and can hardly estimate the LD.

Thus, regardless of the transmission technology (i.e. CoD or DD), *signal overlap* is a critical issue for FON. However, the *misalignment* between the signal and the receiver filter happens only in DD systems.

Therefore, cost-effective approaches to monitor FONs are needed to allow network operators to take prompt actions in case of improper operation of a device in their domain. Most of the current surveillance systems rely on the capabilities of coherent receivers to collect measurements [Do15]. Regardless of their complexity, these approaches limit the performance of surveillance systems, which are intended to monitor the whole domain in real-time with the minimum extra cost and complexity. Moreover, such measurements are not available if DD systems are exploited.

Similarly as for the detection of soft-failures in SSONs, the availability of OSAs deployable in the optical nodes opens a new horizon for the development of surveillance platforms that can benefit from the analysis of optical spectrum. In this regard, optical spectrum monitoring, with the help of cost-effective OSA, can be considered as a novel solution for monitoring the proper operation of lightpaths in FONs.

One of the key parameters determining the cost of the OSAs is their resolution, which has a significant impact on the accuracy of optical spectrum-based detection and identification solutions. Therefore, a target OSA resolution should be studied. It is worth mentioning that while traffic broadcasting is seen as a drawback of FONs; it allows using one single OSA installed in the last span to acquire all signals in the FON. Therefore, remarkable cost savings can be achieved by properly delimiting their target resolution.

Considering the above-mentioned challenges of FONs and the benefits brought about by monitoring optical spectrum using OSAs, it is of paramount importance to

devise robust and reliable monitoring solutions that allow FON to exploit novel CoD and DD based transmission systems while go into operation with lower margins.

3.3 Autonomic Transmission Exploiting SOP

An autonomic system is defined as a system capable of performing self-management, including self-configuration, self-optimization, and self-protection [RFC7575]. Considering that definition, automating the optical transport layer is a key enabler for the sustainable growth of the networks, due in part to the increase in the dynamicity and complexity of the networks. In fact, network-wide performance can be substantially enhanced while reducing the margins and minimizing the human intervention.

Development of autonomic networking requires programmable infrastructure, analytics and intelligence, and software defined control [Ciena]. Autonomic networking in the optical layer has received significant attention in the recent years to cope with the future coming heterogeneous services with bursty traffic. One of the main steps in the realization of such paradigm is monitoring to make every aspect of the network transparent to a central controller. Then, depending on the type of monitoring data, further step can be followed to analyze the data and extract meaningful clues to act.

Even though there is variety of parameters to monitor [Do15], SOP has received little attention for network automation. It is possible to monitor the SOP traces either with the help of DSP units of coherent receivers [Bo18] in the end-nodes or following a device-based technique [Si17]. In [Si17], it has been shown that SOP can be considered as an indicator to reveal fiber disturbance as a sign of fiber break. For instance, ground digging near a fiber may shake a fiber cable well before breaking it [Si17]. In such occasion, the central controller could perform proactive rerouting prior to the fiber break avoiding connection disruption. In [Bo17], it has been demonstrated that SOP monitoring even allow classifying the type of vibration the fiber experiences. What is missing in the literature is to find a correlation between pre-FEC BER and the evolution of SOP traces over time. Considering that such solution can be developed, we could predict the evolution of BER by monitoring the SOP traces.

3.4 SDM Switching Technologies

The implementation of switching solutions for the switching paradigms described before requires the design of new SDM switching nodes [Ma15.1], [Fe15]. Ind-Sw can be realized using node architecture such as the one shown in Fig. 3-1 for a

route-and-select ROADM configuration, applicable for transmission fiber without mixing/coupling. It is composed of some conventional WSSs, one per spatial dimension, degree and ingress/egress port. Commercially available 1×5, 1×9 or 1×20 WSSs can be employed since the port count is not a limiting factor in this case.

The selection of one or another realization will depend on a number of factors, such as the required ROADM operation (e.g. whether "spatial lane change" —i.e. rerouting from one spatial dimension (fiber/core/mode) to another [Ma15.1]— are allowed) and the nodal degree —i.e. the number of available directions— of the network nodes. Fig. 3-1 depicts a node architecture enabling Ind-Sw without spatial lane change for an SDM network with $S = 4$ spatial dimensions and a node with degree $D = 3$. A/D modules (not shown in Fig. 3-1), depending on the chosen A/D module, can allow for several degrees of operational flexibility: they can be colorless, directionless and contentionless. Colorless, contentionless ROADM architectures were proposed in [Ma15.1].

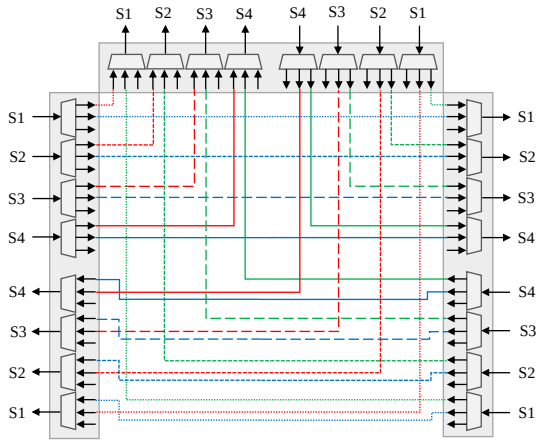


Fig. 3-1. SDM node architecture, for S=4 and D=3, enabling Ind-Sw

Colorless, directionless, contentionless ROADM architectures are based on Multicast Switches (MCS) [Ar15], [Wa11] or M×N WSSs [Ma17.1]. A free-space port-reconfigurable WSS serving as a 7×21 switch was reported in [Pa15]. An 8×28 WSS combining waveguide and free- space optics was demonstrated in [Ik16.1], [Ik16.2]. Assuming Colorless, Directionless and Contentionless (CDC) operation, so that the node architecture depicted in Fig. 3-1 requires two (or four, to avoid a single point of failure) MCSs or M×N WSSs with 12 input ports and as many output ports as the number of Tp that need aggregating. To relax the requirement of High-Port-Count (HPC) A/D modules, K adds/drop modules with port count M×N/K can be used instead, in which case K more output ports per ingress/egress WSS are required.

FrJ-Sw and J-Sw make necessary a redesign of the WSSs. SG WSSs compatible with BuSMF, FMFs or MCFs were reported in [Ne14], [Ca14], [Fo15]. They are

configured to operate as $S \times (I \times O)$ WSSs, i.e. they direct I input ports, each carrying S spatial modes/cores, toward O output ports using spatial diversity. In [Ne14], a commercial 1×20 WSS was reconfigured to implement a 7×(1×2) WSS for a 7-core MCF with spatial diversity. In [Fo15], the WSS port count was increased to 57 by replacing the input/output optical frontend with a two-dimensional SMF array laid out in a 19×3 grid. The operation of this HPC WSS as a 6×(1×8) SDM WSS was demonstrated in a heterogeneous SDM network with BuSMFs, FMFs and MCFs. Spatial diversity solutions in support of FMFs, MCFs or FM-MCFs, introduce interfaces capable of converting the MCF/FMF inputs/outputs into BuSMFs, such as photonic lanterns for FMFs [Hu15] and FM-MCF [Mi14], [Sa16.2], or MCF fan-ins/fan-outs based on 3D direct laser writing [Th07] or tapered fiber bundles [Zh10] for MCFs, need to be interposed between the fiber and the WSS.

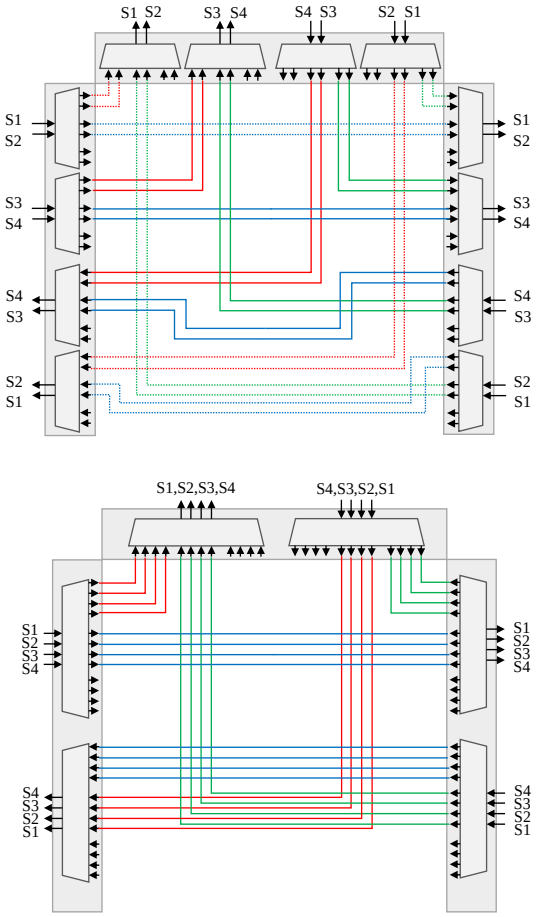


Fig. 3-2. SDM node architecture enabling a) FrJ-Sw and b) J-Sw

An alternative WSS modification in support of joint switching applicable for the case of FMF transmission is to replace the WSS input/output SMF array with an FMF array. This requires a modification of the anamorphic beam-expansion optics

so that the spectral resolving power is increased, thus mitigating the impact of channel bandwidth reduction due to mode dependent loss at channel edges [Ma15.2]. FMF-WSSs are allowing joint switching of three spatial modes in a 3×(1×9) and 3×(1×2) configuration were demonstrated in [Ma15.2] and [Gu13], respectively.

By making use of SG WSSs, the FrJ-Sw and J-Sw paradigms enable reducing the number of necessary WSSs to $2\lceil S/G \rceil$ and 2, respectively, per degree, as illustrated in Fig. 3-2(b) and (c), but the required port count increases by a factor of G and S, respectively. In the case of a large number of spatial dimensions, the joint-switching paradigm makes it necessary that new realizations be explored. Recently, a technique to increase the port count by implementing a non-trivial unitary transform between an array of Gaussian beams and an array of overlapping Sinc functions (resulting in rectangular shaped beams on the LCoS) has proved to be capable of doubling the WSS port count compared to standard Gaussian illumination [Fo16].

3.5 Resource Allocation in SDM Networks

Resource allocation is a well-known problem in optical networks and has been widely investigated for WDM networks and EONs. Even though alternative algorithms/solutions have been proposed so far addressing this problem, the core idea in all of them is to find an available wavelength (in a WDM network) or a contiguous portion of available spectrum (in an EON) on a set of consecutive links, connecting two end-nodes, to establish a lightpath between them. However, there are some technological/strategical constraints, which must be taken into account while proposing a resource allocation option. For instance, the same wavelength or the same portion of contiguous spectrum must be available on all the links connecting two end-nodes to avoid placing costly wavelength converters in the intermediate nodes where optical bypass operations take place. This constraint is known as *wavelength/spectrum continuity*, which is the only one applying to route and wavelength assignment (RWA), the conventional resource allocation problem in WDM networks.

By emerging an extra degree of freedom introduced in EONs (i.e. the choice of variable channel spacing), an extra constraint has been added to RWA problem. To address the RSA, the popular resource allocation problem in EONs, in addition to spectrum continuity, *spectrum contiguity* must be considered. Spectrum contiguity means that, if a certain connection requires more than one frequency slots, these frequency slots must be adjacent on the optical spectrum. Modulation format is another degree of freedom exploited in EONs due to the capability provided by bandwidth-variable TRxs. Thus, RSA problem has been extended to consider modulation format, and it becomes Routing, Modulation format, and Spectrum

Assignment (RMSA) problem. RMSA problems are typically optimized with the aim of minimizing spectrum utilization, energy consumption, or Opex.

With the introduction of space dimension, RMSA problems must be upgraded to take into account space domain. Therefore, RMSA problem becomes Routing, Space, Modulation format, and Spectrum Assignment (RSMSA) in the context of SDM based optical networks. In this section, we try to identify the potential capabilities/restrictions brought about by the space dimension. In particular, we discuss the extra constraints imposed by enabling technologies of SDM (i.e. fibers and switches), which affect the RSMSA problem. Ultimately, several channel allocation options will be presented, and their benefits/drawbacks will be discussed.

The two main technology areas limiting channel allocation and routing options in an SDM based optical network are *fiber type* and *switching paradigms*. SDM fibers affect the resource allocation options through imposing various physical impairments to spectrally/spatially multiplexed channels (i.e. superchannels) while switching paradigms do so by determining the routing properties of the multiplexed channel at the SDM nodes.

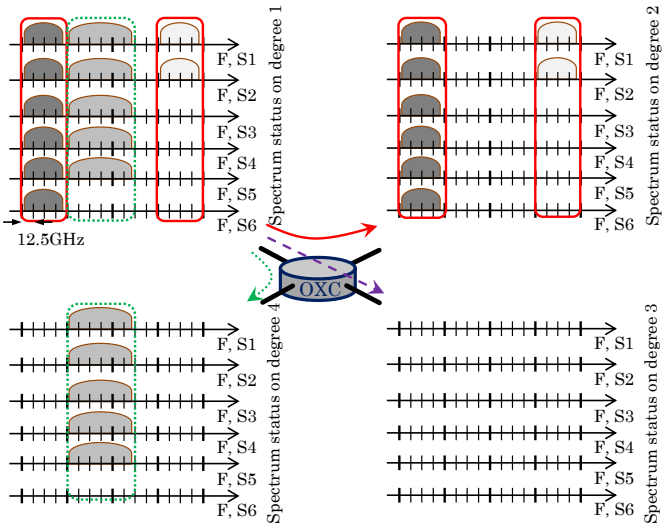


Fig. 3-3. Routing principle in the case of J-Sw scheme

Fig. 3-3 - Fig. 3-7 show the routing options in a 4-degree SDM node utilizing alternative SDM switching paradigms assuming that an SDM fiber with six spatial channels is in place. To illustrate how three SDM switching paradigms determine the routing operation, some exemplary occupied spectral and spatial slots (each one of those dark, grey, and white rectangle-like shapes denotes one spectral-spatial slot) are considered on degree 1 of the depicted node of Fig. 3-3 - Fig. 3-7 .Note, even though the same spectrum status is assumed on degree 1 of all options, they are distributed among a different number of demands for different switching paradigms. The spectral-spatial slots bounded by solid lines, dashed lines, and dotted lines are routed together from degree 1 to degree 2, 3, and 4, respectively.

Each of these rectangles denotes an independent demand required to be routed separately.

As is shown in Fig. 3-3, a certain portion of the spectrum on all spatial dimensions must be switched together when J-Sw is used. Therefore, if demand is small and it only spreads over some of the spatial dimensions, the available spectrum on the rest of the spatial dimensions cannot be allocated to other demands and remains unutilized due to the coarse spatial granularity of J-Sw.

Utilizing switching paradigms with finer spatial granularity relaxes this constraint if we are not limited by the physical impairment of the fibers. Routing of spectral-spatial slots assuming FrJ-Sw with G=3 is presented in Fig. 3-4. The value of G equal to 3 means that the six spatial dimensions are divided into two groups of size 3 and each group can be freely switched from/to any degree of the ROADM. This feature allows the allocation of distinctive demands on individual groups. However, if a single demand is big enough to be spread over all spatial dimensions, FrJ-Sw routes all dimensions together as a single entity, similar to the case of J-Sw.

In Fig. 3-4 and Fig. 3-5, the spectral-spatial slots are kept on the same portion of the optical spectrum and the same spatial dimensions while routing to another degree. It means that, in addition to the *spectrum continuity* constraint, *space continuity* is considered. Space continuity is a constraint, which must be ensured while devising RSMSA algorithms for SDM networks based on coupled transmission media. Space continuity constraint is granted in the architecture of J-Sw and FrJ-Sw shown in Fig. 3-4 and Fig. 3-5, respectively. This constraint is relaxed for fibers having uncoupled spatial groups (e.g. FM-MCFs) or uncoupled individual spatial dimensions (e.g. weakly-coupled MCFs), thus allowing LC operation at switching nodes. It is worth noting that, while LC operation is allowed between groups for fibers with uncoupled spatial groups, space continuity must yet be considered inside each group.

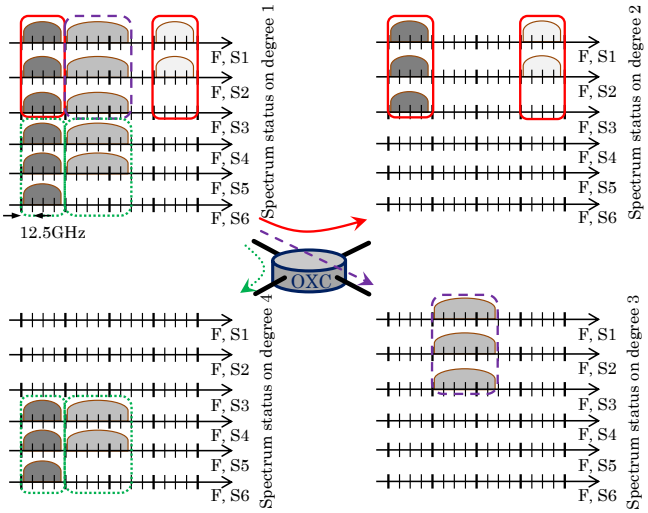


Fig. 3-4. Routing principle for FrJ-Sw scheme (G=3) without LC support

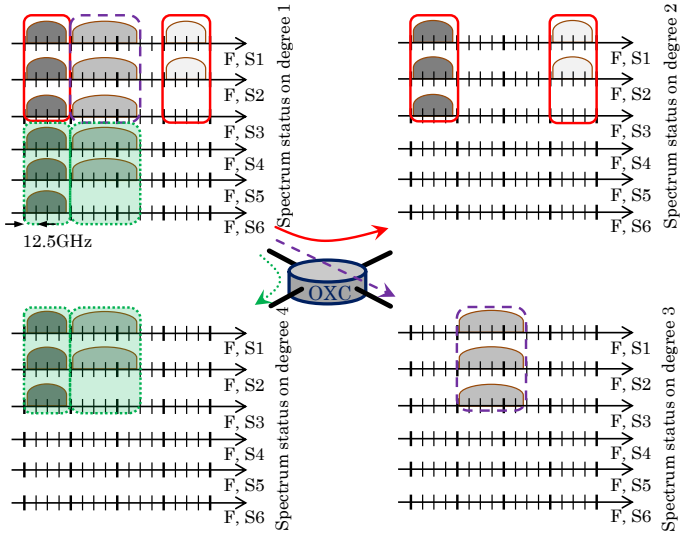


Fig. 3-5. Routing principle for FrJ-Sw scheme (G=3) with LC support

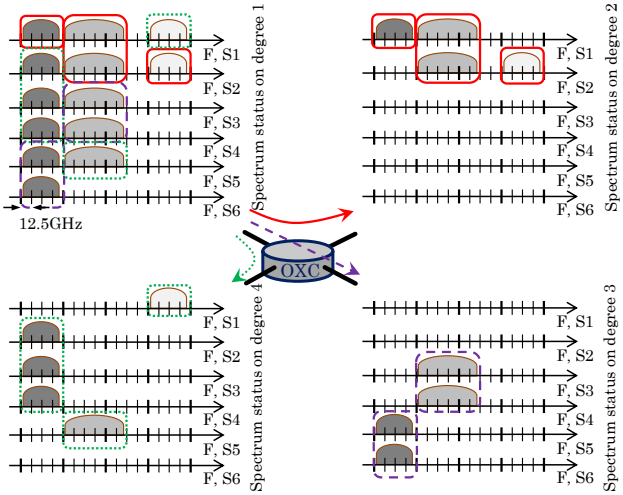


Fig. 3-6. Routing principle for Ind-Sw scheme without LC support

FrJ-Sw with LC allows the routing of spatial groups independently to any group of spatial dimensions on any direction of the ROADM as shown in Fig. 3-5. Rectangles drawn with dashed-lines show the spectral-spatial slots, which undergo LC operation. The two filled bounded slots by dotted lines illustrated in Fig. 3-5 are routed from one spatial group on degree 1 to another spatial group on degree 4. This property allows the implementation of more flexible/advanced RSMSA algorithms.

Finally, routing of spectral-spatial slots exploiting Ind-Sw is presented in Fig. 3-6 and Fig. 3-7. In Fig. 3-6 spectral-spatial slots can be freely routed to any direction, but space continuity constraint is applied. Fig. 3-7 shows the most flexible routing scheme in which the space continuity constraint is relaxed and any spectral-spatial slot can be independently directed to any spatial dimension on any degree of the

ROADM. Ind-Sw with LC can be exploited for SDM networks based on weakly-coupled MCF or BuSMFs, and it must be avoided for SDM networks based on other fiber types (e.g. FM-MCF, FMF). The capability of performing LC increases the complexity and thus the implementation cost of the SDM node (remember that (a) shows only the architecture of an SDM node based on Ind-Sw without LC). Note that wavelength conversion is not allowed in any of these options and, therefore, spectrum continuity must always be ensured.

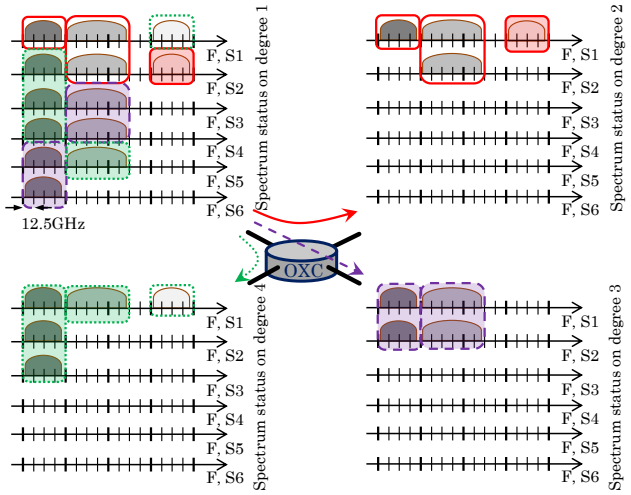
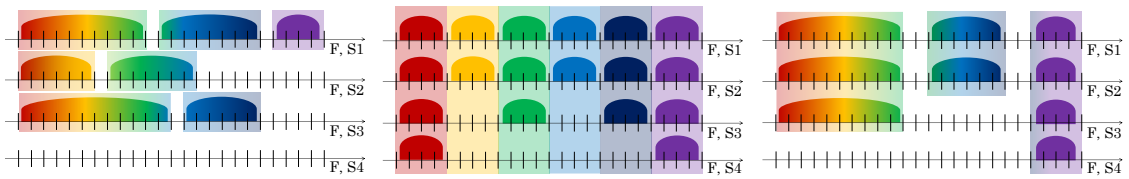


Fig. 3-7. Routing principle for Ind-Sw scheme with LC support

As previously discussed, RSMSA algorithms depend on the enabling technologies, mainly the fiber type and the switching paradigm, on which an SDM network is deployed. Ind-Sw with LC provides the highest level of flexibility for RSMSA algorithms. However, it requires the most complex and costly node architecture. At the opposite extreme, J-Sw offers the lowest level of flexibility for RSMSA algorithms and the simplest and most cost-effective node architecture. Even though some switching paradigms suffer from lower routing flexibility, they perform well in some specific cases. We discuss further this issue in the next two sections where extensive comparative simulation results are presented evaluating SDM switching paradigm in various cases.

Even though the fiber types and switching paradigms are two main technology areas determining the channel allocation properties, spectral/spatial multiplexing techniques, forming end-to-end optical transport channels, are defined by the TRx technology. Three channel allocation options, which are followed by three approaches that a TRx can form an optical transport channel, are presented in Fig. 3-8. Fig. 3-8(a) shows the case in which demands are transported in the form of spectral Sp-Chs, i.e. demands spread over a contiguous portion of the spectrum, enough to accommodate them, in a single spatial dimension. Spectral Sp-Chs can improve the spectral efficiency by decreasing the guard bands between subchannels.



(a) Spectral Sp-Ch (b) Spatial Sp-Ch (c) Spectral-Spatial Sp-Ch

Fig. 3-8. Three resource allocation policies for SDM networks

Nonetheless, SDM networks based on spectral Sp-Ch require Ind-Sw based switching nodes, which are the most costly switching elements. Practically speaking, spectral Sp-Ch allocation policy can be used only for SDM networks based on uncoupled/weakly-coupled transmission media. On the other hand, Fig. 3-8(b) illustrates the case in which demands are transported in form of spatial Sp-Chs, i.e. demands spread across a number or all of the spatial dimensions over a given spectral slice (e.g. one optical carrier). Spatial Sp-ch allocation policy can be utilizied with any type of transmission medium, but is necessary for SDM networks based on coupled transmission media. Spatial Sp-Ch allocation policy allows: i) decreasing the optical switching complexity, since spatial channels are switched in groups (i.e. Sp-Chs) rather than independently —at the price, of course, of a potential reduction in routing flexibility— [Fe15] ; and (ii) SDM networks based on spatial Sp-Ch allocation can benefit from additional cost reduction due to the possibility of sharing network elements among different spatial dimensions (e.g., a number of Sp-Ch constituents can share lasers and DSP modules, which can lead to cost and power consumption savings of integrated transceivers in SDM networks. A third possible policy is spectral-spatial Sp-Ch, shown in Fig. 3-8(c), would emerge from the combination of the aforementioned spectral and spatial Sp-Ch allocation policies, i.e. demands spread across a number or all of the spatial dimensions over a contiguous portion of spectrum. One can say that spatial Sp-Ch is a particular case of spectral-spatial Sp-Ch when the spectral width of spectral-spatial Sp-Chs is as narrow as a single optical carrier.

Space dimension provides a huge capacity and flexibility to allocate spectral and spatial channels. However, in order to propose an RSMSA algorithm that efficiently utilizes spectral and spatial resources, we must take into account the physical charateristics of transmission media, the routing constraints of switching paradigms, and the capabilities of the TRxs.

Several works have dealt with networking level performance evaluations of SDM networks and, in particular, resource allocation strategies. In [Mu14], a Routing, Spectrum and Core Assignment (RSCA) algorithm was proposed aiming at minimizing the spectrum utilization in a weakly-coupled MCF-based SDM network. The problem was formulated as an integer linear programming, and a heuristic algorithm was proposed. In [Mu15], considering the spectral Sp-Ch allocation policy, the authors extended the RSCA algorithm presented in [Mu14] to

include the modulation format and jointly optimize the switching and spectrum resource efficiency in weakly-coupled MCF-based SDM networks implemented using architecture on demand and static ROADMs. In [To15], routing, spectrum, core, and mode assignment methods were proposed, which exploited prioritized area concept and XT awareness depending on whether or not MCF or MMF were affected by inter-core/inter-mode XT. Note that the spatial Sp-Ch allocation policy was not investigated in any of these works. In [Ro17], exploiting the spatial Sp-Ch allocation policy, the routing, modulation format, baud rate, and mode assignment problem was addressed with the aim of showing the benefits of FMFs in metro networks. Note that, the two main technology areas limiting channel allocation and routing options in an SDM based optical network are fiber type and switching paradigms. However, a comprehensive performance evaluation of SDM networks realizing over different types of switching paradigm and transmission media is missing from the literature.

3.6 **Conclusions**

In this chapter, we have reviewed the state-of-the-art of relevant works related to the goals of this PhD thesis. Table 3-1 summarizes the reviewed work upon which we build the cornerstone of the studies of this PhD thesis.

Table 3-1: State-of-the-art summary

Goals	References
Optical Spectrum Monitoring	[Finisar.1] [Finisar.2] [Ki04] [Ko16] [Ma05] [Po17] [Pu11] [Ra18.1] [APV17.1] [APV17.2] [Za13]
Benefits of SOP Tracking for Autonomic Transmission	[Ciena] [Bo17] [Do15] [Gi18] [Ha09] [Ku17] [Po17] [Ra18.1] [RFC7575] [Sa12] [Si17] [Ve18.1]
SDM Network Performance Evaluation	[An16] [Ar15] [Dh16] [Fe12] [Fu14] [Ji16] [Kl15] [Ma15.1] [Ma17.1] [Ma17.2] [Mu14] [Mu15] [Pa15] [Pr15] [Ra16.1] [Ra13] [Ro16] [Ro17] [Sa15] [To17] [Wi12] [Wi14]

The next chapters present the contributions of this PhD Thesis to fulfill the goals described in Chapter 1.

Chapter 4

Feature-based Spectrum Monitoring

In this chapter, we start exploring the benefits of analyzing the optical spectrum of lightpaths for soft-failure detection and identification in Spectrum Switched Optical Network (SSON). We present a framework exploiting ML-based algorithms that uses descriptive models of the optical spectrum of a lightpath in different points along its route to detect whether the optical signal experiences anomalies reflecting a failure in the intermediate WSSs. Our proposal targets the two most common filter-related soft-failures; Filter Shift (FS) and Filter Tightening (FT), which noticeably deform the expected shape of the optical spectrum. In this regard, filter cascading is a key challenge as it affects the shape of the optical spectrum similarly to FT. The approaches are specifically designed to avoid the misclassification of properly operating signals when normal filter cascading effects are present. DT and SVM algorithms are considered as candidate ML algorithms to perform the classification task. Extensive numerical results are ultimately presented to compare the performance of the proposed approaches in terms of accuracy and robustness.

4.1 Motivation and Objectives

Failure identification and localization can reduce failure repair times greatly. As described in Section 3.1, there have been several works targeting monitoring the performance of lightpaths at the Tp side to verify their proper operation, as well as to detect BER degradations thus, anticipating connection disruptions. However, monitoring the signal at the egress node does not allow localizing failures and therefore, monitoring techniques to analyze and evaluate QoT in-line are required. In this regard, the availability of cost-effective OSAs, integratable in the optical nodes, allows real-time monitoring of the optical spectra of the lightpaths.

Therefore, optical spectrum features can be exploited by ML-based algorithms to detect degradations and identify failures.

In this chapter, we study two different approaches to detect filter related failures. The approaches are based on a set of classifiers that make predictions using meaningful features extracted from the optical spectrum. These approaches can be considered to deal with filter cascading effects: *i*) the *multi-classifier approach*, in which different classifiers are employed for signals experiencing different levels of filter cascading and *ii)* the *single-classifier approach,* in which the lightpaths' features are pre-processed to compensate for the filter cascading effect allowing the use of a single classifier for lightpaths disregarding the level of filter cascading. Ultimately, the optical spectrum analysis can be used by centralized algorithms able to localize failures in the network.

The proposed approaches for failure detection and identification can be deployed in the MDA agents, close to the devices generating measurements, whereas other algorithms, including the one for failure localization, need to be deployed in the MDA controller, so as to provide the global network vision required.

4.2 OSA for Soft-Failure Detection and Identification

Real-time optical spectrum monitoring provides opportunities for soft-failure detection and identification; particularly, those failures significantly deforming the optical spectrum of a lightpath. For precise detection and identification, algorithms need to be capable of classifying a properly operating lightpath from a failed one, which entails that a set of descriptive features should be identified for classification purposes. Building upon such features, ML-based classifiers can be trained to perform the classification task.

4.2.1 Soft-failure Detection, Identification, and Localization

The failure detection, identification, and localization process involves modules running in the MDA agents and modules running in the MDA controller, as shown in Fig. 4-1. In the MDA controller, the FailurE causE Localization for optIcal NetworkinG (FEELING) algorithm is primarily responsible for supervising the failure detection and identification modules running in the MDA agents. Ultimately, it performs the failure localization task.

Optical spectra acquired by OSAs are collected and becomes available in MDA agents, where it is used to feed the Feature Extraction (FeX) module. An example of a 30-GBaud QPSK modulated optical spectrum acquired by an OSA of 312.5 MHz resolution is shown in Fig. 4-2.

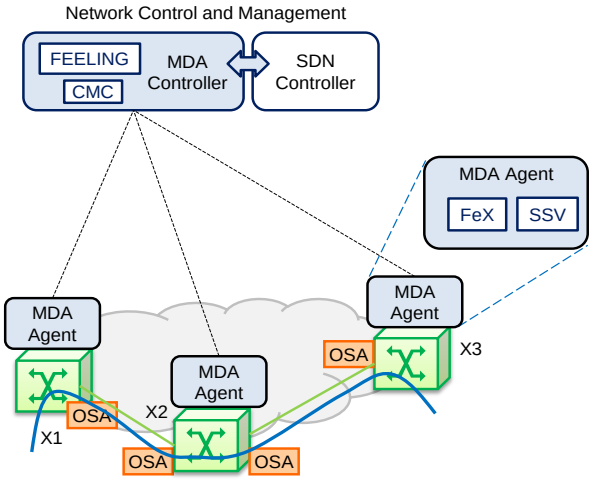


Fig. 4-1. Architecture for failure identification and localization

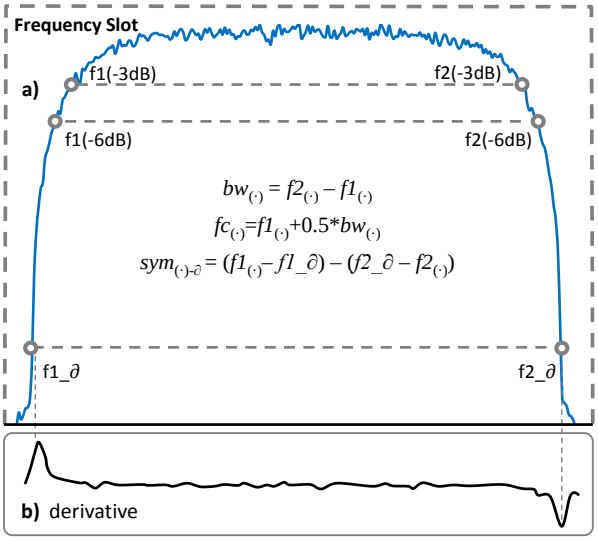


Fig. 4-2. Relevant points of a QPSK modulated signal

The FeX module processes the acquired optical spectrum for a given lightpath, which consists of an ordered list of frequency-power (<*f*, *p*>) pairs in the allocated frequency slot. After equalizing power, so the maximum power is set to be 0 dBm, the derivative of the power with respect to the frequency is computed; Fig. 4-2(b) illustrates the derivative of the example optical signal, where sharp convexity can be observed close to the edges. Next, the FeX module characterizes the mean (μ) and the standard deviation (σ) of the power around the central frequency ($fc \pm \Delta f$), as well as a set of primary features computed as cut-off points of the signal with the following power levels:

- edges of the signal computed using the derivative, denoted as ∂,
- a family of power levels computed w.r.t. μ-$k\sigma$, denoted as $k\sigma$, and

- a family of power levels computed with respect to μ-*mdB*, denoted as -*mdB*. Each of these power levels generates a couple of cut-off points denoted as $f1_{(\cdot)}$ and $f2_{(\cdot)}$.

In addition, the assigned frequency slot is denoted as $f1_{\text{slot}}$, $f2_{\text{slot}}$. Other features, which are computed as linear combinations of the relevant points, focus on characterizing a given optical signal (see embedded equations in Fig. 4-2(a)); they include: bandwidth (*bw*), central frequency (*fc*), and symmetry (*sym*) with respect to a reference (frequency slot or derivatives).

Some features are more appropriate for filter-related failure detection and identification, such as bandwidth and symmetry, whereas other features, such as the central frequency, are more appropriate for laser drift identification. For illustrative purposes, Fig. 4-3 shows how some of the identified features evolve with the severity levels of FS and FT. Fig. 4-3(a, c) show the evolution of $bw_{\text{-3dB}}$, $bw_{\text{-6dB}}$, $sym_{\text{-3dB-}\partial}$, and $sym_{\text{-6dB-}\partial}$ w.r.t the magnitude of FS; while Fig. 4-3(b, d) show the evolution of those features w.r.t magnitude of FT. It can be interpreted that $sym_{\text{-3dB-}\partial}$ and $sym_{\text{-6dB-}\partial}$ are meaningful features to identify FS, while $bw_{\text{-3dB}}$ and $bw_{\text{-6dB}}$ are more sensitive to the effects of FT.

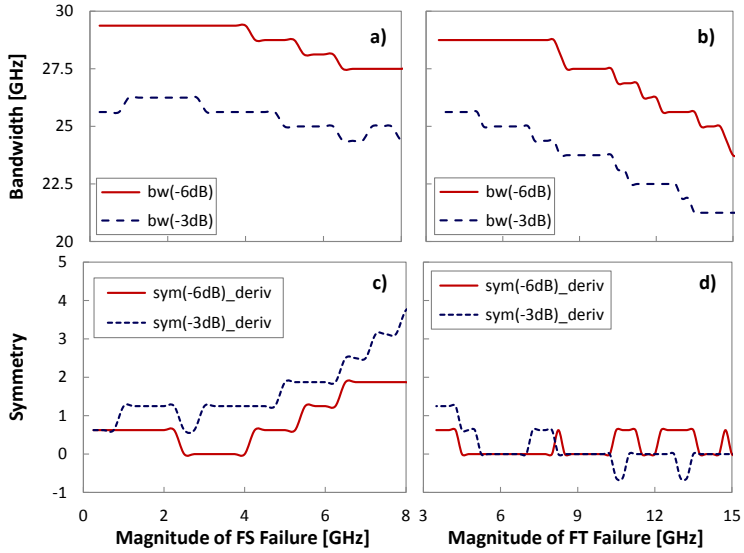


Fig. 4-3. The evolution of the features w.r.t magnitude of FS (a, c) and FT (b, d)

When the extracted features from the measured signal are available, a classification module, named Signal Spectrum Verification (SSV), running also in the MDA agents analyzes them to detect a soft-failure. The SSV module is implemented as a multiclass classifier that produces a diagnosis, which consists of: *i*) a predicted class among '*Normal*', and '*FilterFailure*'; and *ii*) a subset of relevant signal points for the predicted class. In the case that a filter failure is detected, another classifier is used to predict whether the failure is due to FS or to FT.

As discussed before, one of the key challenges in the identification of filter related failures is the misclassification of a normal signal that has passed through several filters, i.e., affected by filter cascading, as a signal that has suffered from filter failure. Therefore, to improve failure identification accuracy, the FEELING algorithm must be able to distinguish between actual failures and normal effects arising from filter cascading. In the next section, we propose and study two approaches to prevent such misclassification.

4.2.2 Options for Classification using FeX

Let us first explain how the FeX relevant features can be used to classify different types of spectra. Fig. 4-4 shows how even a pair of such features can discriminate different types of spectra. Assuming a set of measurements after 2 WSSs and belonging to normal operation, FS, and FT, Fig. 4-4 (a-b) show $sym_{\text{-3dB-}\partial}$ w.r.t $bw_{\text{-3dB}}$ and $sym_{\text{-6dB-}\partial}$., respectively. As represented, observations belonging to different classes can be easily discriminated with even just these two features in place. Now, let's take the same set of observations after 12 WSSs; the results are plotted in Fig. 4-4 (c-d). As shown, it becomes very challenging to distinguish the observations belonging to different classes. It is worth mentioning that as a result of filter cascading signal features change in a similar way it happens when a FT failure takes place.

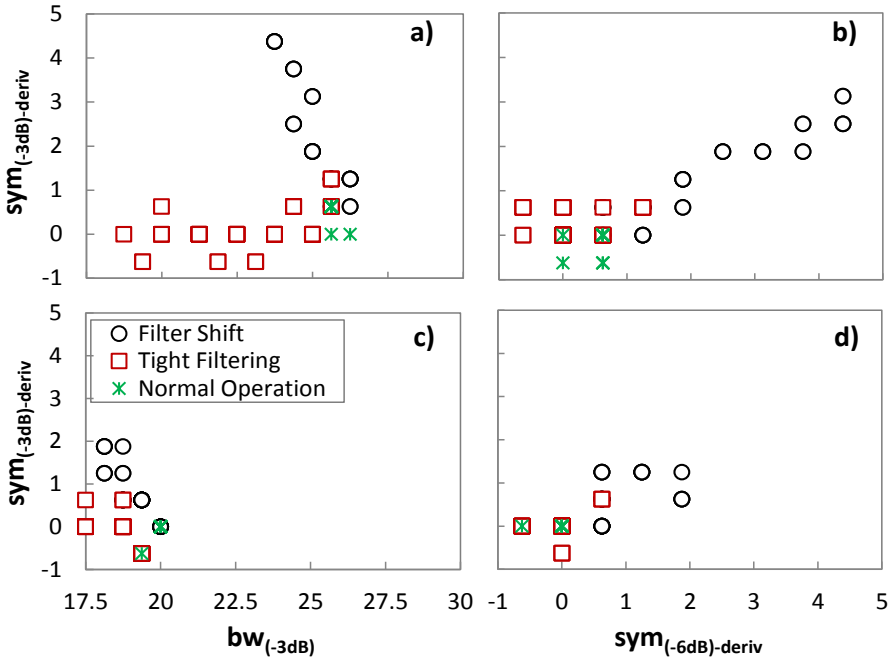


Fig. 4-4. The efficiency of the identified features for classification after 2 WSSs (a, b) and after 4 WSSs (c, d)

Let's take a closer look at the figures to understand this issue. As shown in Fig. 4-4 (a-b), just the observations belonging to FT are gathered in the bottom left corner of

the figures; however, in Fig. 4-4 (c-d), almost all the observations are gathered in the bottom left corner of the figures. In other words, filter cascading effect push the identified features in the direction that all observations look like FT observations. This increases the likelihood of misclassifying a properly operating lightpath as a failed one. In the following, we propose two different strategies preventing such misclassification. The strategies, built in the SSV module, are based on processing the features extracted by the FeX module. Selected features for classification are: bw_{∂}, bw_{5o}, $bw\text{-}_{3dB}$, $bw\text{-}_{6dB}$, $sym_{5o\text{-}\partial}$, $sym_{\text{-}3dB\text{-}\partial}$, *and* $sym_{\text{-}6dB\text{-}\partial}$.

4.2.2.1 Multi-Classifier Approach

The most straightforward solution is to use different classifiers as a function of the number of WSSs that a given lightpath has passed through. As shown in Fig. 4-5(a), a set of classifiers are required in every intermediate node and the appropriate one is used when an optical spectrum is acquired. This approach can be considered as the baseline, as the selected classifier decides based on the features extracted directly from the acquired spectrum and do not need any kind of feature pre-processing. However, a very large dataset of optical spectra with different levels of filter cascading is required for training all the classifiers, which is the main drawback of this approach.

To avoid using multiple classifiers, some pre-processing needs to be done so one single classifier can be used despite the level of filter cascading. The second approach proposes a strategy to pre-process the extracted features.

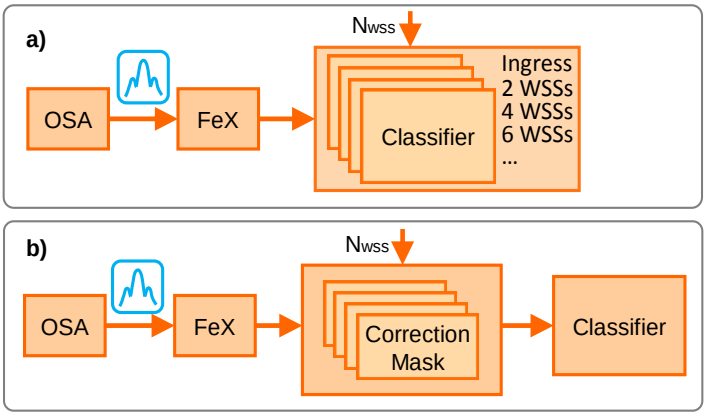


Fig. 4-5. Feature-based approaches: a) multi-classifier, b) single-classifier.

4.2.2.2 Single-Classifier Approach

Filter cascading strongly affect some of the features that a classifier uses for prediction. Therefore, if the alteration of those features due to filter cascading is compensated, a single classifier could be considered regardless of the number of filters a signal passes through. The features of a signal acquired after passing N filters can be compensated by adding/subtracting the differences between the values of a properly configured signal at that node w.r.t. those just after the Tp.

These differences are stored in a vector called *correction mask*; note that, different levels of filter cascading require different correction masks to be used.

Correction masks can be computed a priori, assuming the effects that the spectrum of a normal signal experiences while passing through different number of filters. It is worth mentioning that the calculation of the correction masks requires just the spectrum of a single properly configured lightpath passing through the desired number of filters, from zero to the maximum allowed cascaded filters in a network; this is in contrast to the previous approach, where the training phase requires that spectral data with different failures and with various magnitudes to be captured after every filter up to the maximum allowed number of filters. The Correction Mask Calculator (CMC) module placed in the MDA controller (see Fig. 4-1) is responsible for generating the correction masks to be sent to the MDA agents. Note that all the correction masks need to be available in the MDA agents, so the proper one can be selected. Following this approach, the classifier can be trained based on the observations of a passing through just a single filter, making the training phase less data-hungry by far compared to the previous approach.

Fig. 4-6 summarizes the failure detection and identification work flow followed in both single-classifier and multi-classifier approaches. Once the optical spectrum of a signal has been acquired, the features are extracted. In case of single classifier approach, the features are corrected applying the specific correction mask that corrects filter cascading effects for the number of filters that the signal has traversed. Next, failure analysis can be carried out;

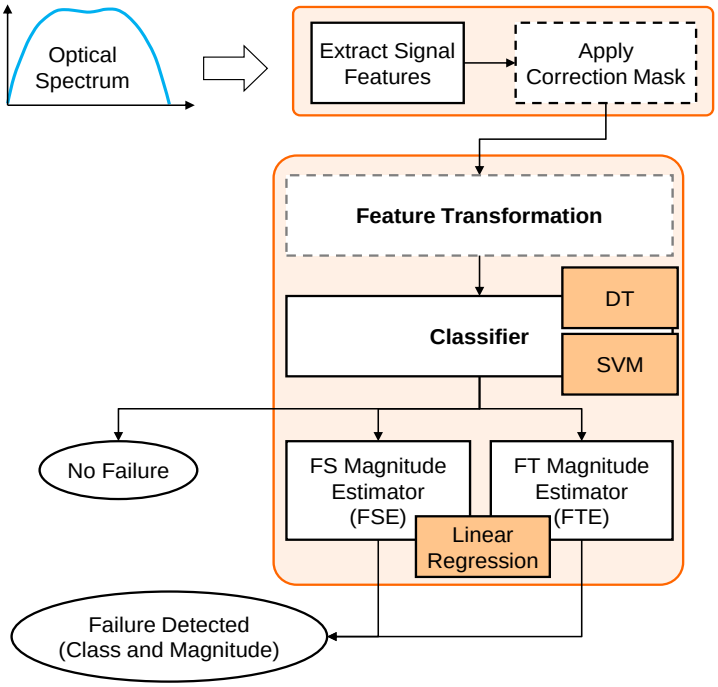


Fig. 4-6. Failure detection and identification workflow

The first alternative classifier is based on DTs, whereas the second one selects SVMs. Both classifiers aim at identifying whether a filter failure is affecting a connection and if so, which is the type of failure: FS or FT. In the case that a failure has been detected, its magnitude needs to be estimated. Two filter failure magnitude estimators can be called depending on the detected failure; both are based on linear regression.

4.2.2.3 Feature Transformation for Single-Classifier Approach

The single classifier approach can be further improved if a feature transformation step is applied as shown in Fig. 4-6. For the sake of simplicity, we consider the magnitude of the failures as additional features for training the classifiers, so we use the magnitude estimator before a failure has been detected. In this way, original features are linearly combined to create new ones that might aggregate information in the hope of improving classification accuracy.

Table 4-1 General classification training algorithm pseudocode

INPUT *dataset*, *Configs*, *maxSplits*
OUTPUT *model*

```
1:  dataset←balanceClassesByReplication(dataset)
2:  for each config in Configs do initialize GoC[config]
3:  for i=1..maxSplits do
4:    <trainingSet,testingSet>←randomSplit(dataset)
5:    initialize configParams
6:    for each config in Configs do
7:      model ← fit(trainingSet, config)
8:      errorTraining ←predict(model, trainingSet)
9:      errorTesting ←predict(model, testingSet)
10:     gocConfig ← GoC[config].addNew()
11:     gocConfig.configParams ← config.params
12:     gocConfig.errorTraining ← errorTraining
13:     gocConfig.errorTesting ← errorTesting
14: bestConfig←computeBestConfig(GoC)
15: return fit(dataset, bestConfig)
```

Let us now get insight about the training process of the classifiers (see pseudocode in Table 4-1). The algorithm receives a dataset of labeled examples that is firstly balanced by adding copies of instances from the class with less number points in the dataset to have the considered classes (normal, FS, FT) equally represented (line 1 in Table 4-1). A set of configurations that contain specific parameters for the classification algorithm selected will be used during the training process.

The parameters considered to fit DTs are the number of observations per leaf; for every n a DT model is obtained. As for SVM fitting, the parameters are the degree of the polynomial kernel (*kernelDegree*) for complexity control and the cost of misclassifying (*misClassCost*) for the size of the SVM. For every configuration, a number of randomly-generated splits of the data set for training and testing will be performed. To store the goodness of each configuration, the *GoC* array will be used in the rest of the algorithm and it is firstly initialized (line 2). Next, a new dataset split is generated; where the training set is used for fitting a model for the classifier with the specific selected configuration (lines 3-7). Once a model is computed, predictions using the training and testing data set are carried out (lines 8-9); the training and testing errors between the model prediction and the actual values are stored in the *GoC* array together with the current configuration parameters (lines 10-13). Finally, the results obtained for the different configurations and training/testing data splits are evaluated to select the configuration with minimum error (line 14). Such configuration is eventually used to fit a model using the whole dataset to improve the algorithm performance (line 15).

4.3 **Illustrative Results**

In this section, we numerically compare the performance of different approaches described in the previous sections. Firstly, the transmission set-up modeled in VPI Photonics is described; the set-up is used to generate the optical spectrum database required for training and testing the proposed algorithms. Next, the two feature-based approaches implemented as DT and SVM classifiers are evaluated thus, revealing their benefits/drawbacks for filter failure identification tasks. Finally, the performance of the three two approaches for failure detection and localization are compared.

4.3.1 VPI Set-up for Data Collection

The VPI set-up is shown in Fig. 4-7(a). In the transmitter side (N1), a 30 GBd DP-QPSK signal is generated, passes through 7 intermediate nodes performing optical switching and ends in a coherent receiver (N9) that compensates for the impairments introduced throughout the transmission. Nodes are interconnected by single mode fiber spans; after each span, optical amplifiers compensate for the accumulated attenuation of the fiber. The transmitter and receiver are assumed to be installed in optical nodes, which are modeled with two 2nd order Gaussian filters emulating optical switching functionality for A/D and pass-through performed by WSSs [Pu11], [Finisar.2]; filters bandwidth is set to 37.5 GHz, leaving 7.5 GHz as a guard band.

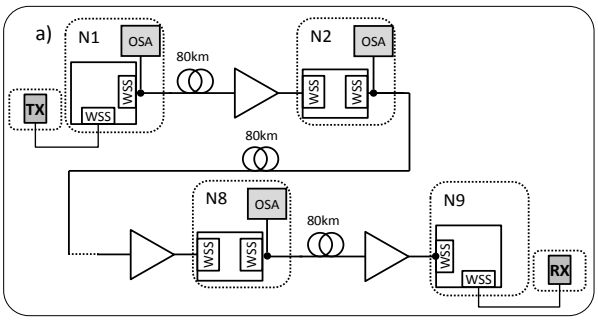


Fig. 4-7. VPI setup

One OSA per outgoing link, configured with 312.5 MHz resolution, is considered in every node to monitor the optical spectrum. As previously discussed, a correction mask should be considered for the features affected by filter cascading, as features get modified while passing through WSSs. Fig. 4-7(b) shows an example of the amount of reduction in the bw_{-3dB} feature of a lightpath in the set-up and the corresponding correction mask, obtained by fitting a 2nd order polynomial.

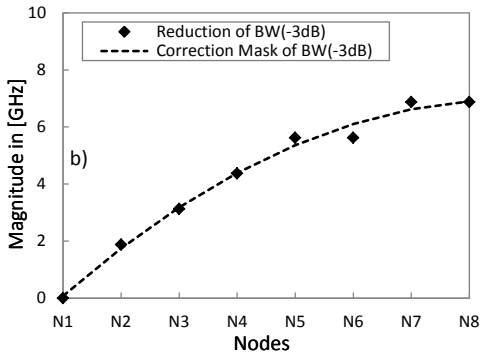


Fig. 4-8. Correction mask of bw_{-3dB} of the setup shown in Fig. 4-7

Aiming at emulating failure scenarios, we modify the characteristics of the 2nd WSS of each node (from N1 to N8) in the set-up; its bandwidth and central frequency are modified to model FT and FS failures, respectively. A large dataset of failures was collected by inducing failures of magnitude in the range [1-8] GHz for FS and in the range [1-15] GHz for FT, both with 0.25 GHz step-size, where the magnitude of FT is defined as the difference between the ideal bandwidth of the filter (37.5 GHz) and its actual bandwidth during the failure.

4.3.2 ML-based Classification Comparison

We compare the performance of the proposed approaches in terms of its *accuracy*, defined as the number of correctly detected failures over the total failures. Fig. 4-10 shows the accuracy of detecting FS and FT, respectively at node N1 in terms of the magnitude of the failure. Note that in N1 both multi-classifier and single-classifier are the same as no filter mask is required.

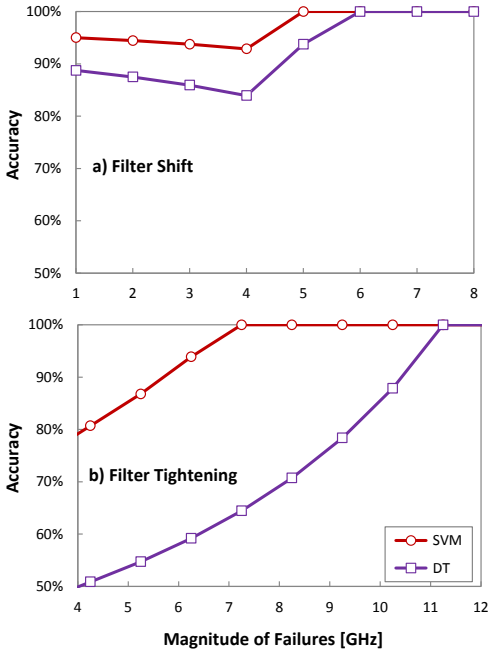


Fig. 4-9. Accuracy at N1 of DT and SVM for FS

Every point in Fig. 4-9 (a-b) aggregates failure scenarios with different magnitude by considering all the observations belonging to a particular failure magnitude and above. As shown, the accuracy of detecting FS larger than 1 GHz is around ~96% when classifiers are based on SVMs, while it hardly approaches 89% when they are based on DTs. On the other hand, the accuracy of SVMs reaches 100% for failures larger than 5 GHz, while this level of accuracy for DTs is achieved for failures larger than 6 GHz. Regarding FT detection, the best accuracy of the proposed classifiers for low magnitudes (below 6 GHz) is around 80% (achieved for SVMs), which is due to the fact that the shape of the optical spectrum is quite similar to the normal scenario, making it very challenging for the classifier to distinguish. This is in contrast to the case of FS, whose effect is more evident even for low magnitudes due to its asymmetric impact on the optical spectrum. For the magnitudes above 7 GHz, the SVM-based classifier perfectly detects the failure. Note that DT-based classifiers achieve perfect accuracy for magnitudes above 10.5 GHz.

Let us now compare both approaches implemented with DT and SVM -based classifiers for detecting failures in all 8 nodes of the set-up. Recall that multiple classifiers are needed for the first approach and several filter masks are required for the second approach. The results are shown in Fig. 4-10(a-b) for FS and FT, respectively, where every point aggregates the results for all the nodes. As observed, SVM-based classifiers significantly outperform DT-based ones in both approaches and failures. As a result, SVM-based classifiers can be selected as the preferred option for feature-based approaches.

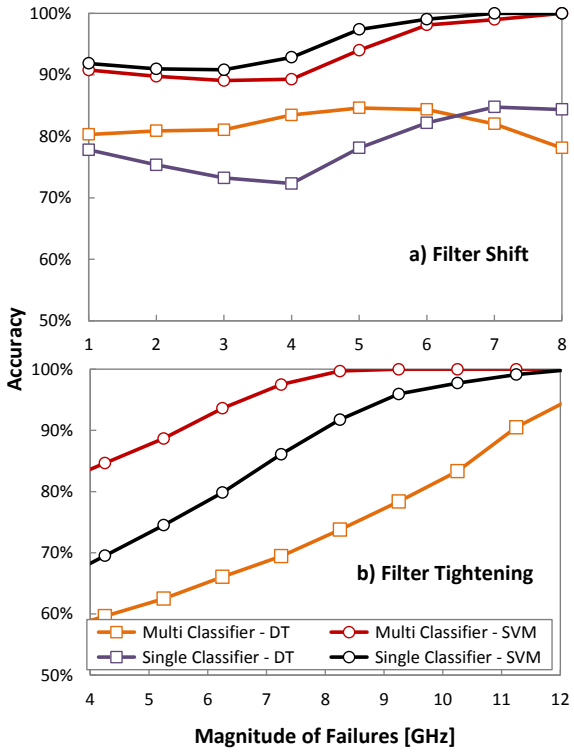


Fig. 4-10. Average accuracy over N1-N8 of DT and SVM for FS (a) and FT (b).

Comparing the different SVM-based approaches, the single-classifier performs slightly better in the case of FS detection while the performance of the multi-classifier approach is much better than the single-classifier one in the case of FT failure. Therefore, we can conclude that training multiple classifiers with the data collected at nodes experiencing different levels of filter cascading performs better than correcting the features with the purpose of using a single classifier, as the impact of filter cascading is similar to the effect of FT on the shape of optical spectrum.

4.3.3 Benefits of Using a Single OSA

In this part, considering the single classifier-based approach, we focus on detecting the failures in some nodes after the point where the failure happens, showcasing the *efficiency* of the proposed methods with respect to the evolution of the optical signal along the transmission path. In addition, by following this approach, the number of utilized OSAs in the network can be reduced. Fig. 4-11 shows the minimum magnitude after which the accuracy of classifiers remains 100% in terms of the location of OSA compared to the point that failures happen; 0 on the x-axis means that OSA is placed at the node where the failure happens (N1 in Fig. 4-7), while 7 means that is placed 7 nodes away from the location of the failure (N8 in Fig. 4-7). It can be understood that the SVM-based classifier is more effective

regardless of the location of the OSA and it perfectly detects the failures above a magnitude threshold. Even though the DT-based classifier shows an acceptable performance for FS failures up to 3 nodes distance from the location of the failure, it fails when considering FT failures.

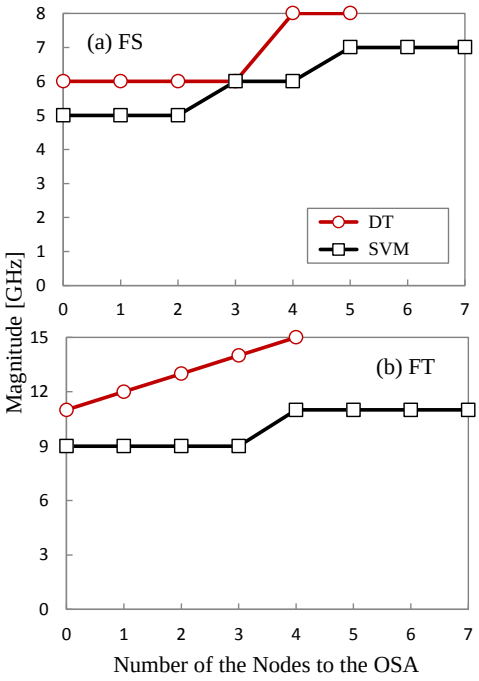


Fig. 4-11. Robustness of the Ml algorithms while one OSA is considered

4.3.4 Benefits of Feature Transformation for Classification

Once the failures are detected, Filter Shift Estimator (FSE) and Filter Tightening Estimator (FTE) can be launched to return the magnitude of the failures (see Fig. 4-6); estimators are based on linear regression. Estimated values of FS and FT with respect to their expected values are illustrated in Fig. 4-12a and Fig. 4-12b, respectively. As shown, the estimators can predict the magnitude of failures with very high accuracy, with Mean Squared Error (MSE) equal to 0.09091 and 0.00583 for FSE and FTE, respectively.

In addition to the use of these estimators to explore the magnitude of the failures, they can be used in the feature transformation step as anticipated in Section 4.2.2.3. In fact, the output of FSE and FTE can be considered as two principal components of an imaginary two-dimensional vector space as shown in Fig. 4-13. In such space, FS and FT failures evolve in different directions of the vector space. As illustrated, the observations belonging to normal operation and the small magnitudes of the failures coincide. However, they become perfectly distinguishable as the magnitude of failures increase.

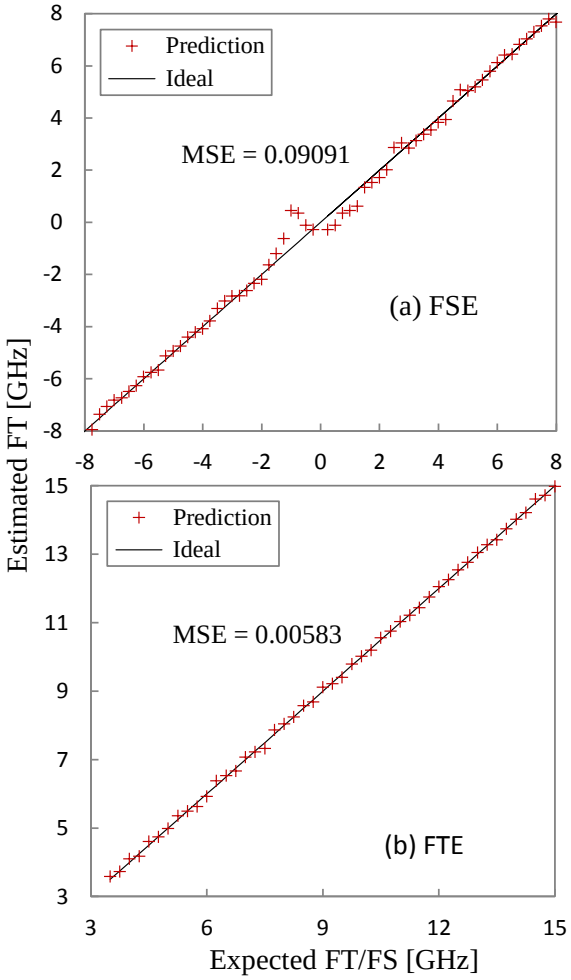


Fig. 4-12. Prediction accuracy of FS (a) and FT (b) estimators.

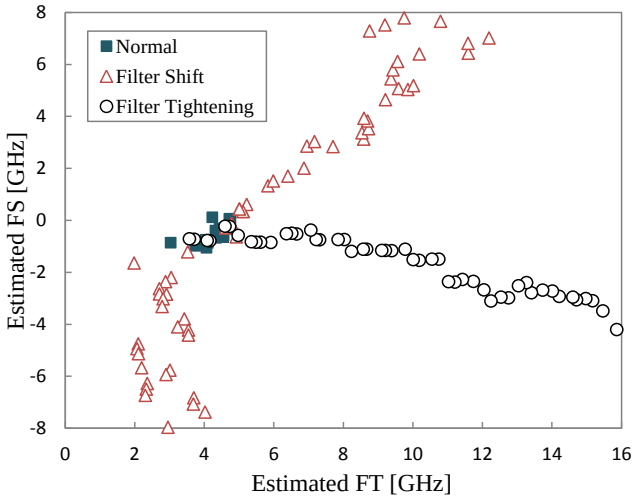


Fig. 4-13. Estimated FS vs estimated FT as a two-dimensional vector space

Let us evaluate the benefits of exploiting the outputs of FSE and FTE estimators as additional features for training the classifiers; Fig. 4-14 presents the obtained results. For the sake of conciseness, we group the magnitudes into three groups of low (L), medium (M), and high (H) magnitudes, instead of reporting all of them independently. Regarding the location of the OSAs, we report just three locations. Analyzing Fig. 4-14(a), one can realize that the accuracy of DTs can be substantially improved, notably for low and medium magnitude, when using the estimations of FSE and FTE as new features. Additionally, it makes the classifier based on DT more robust while using OSAs far away from the location of the failures. However, it yet cannot outperform the classifier based on SVMs, even with these additional features. We also see that adding these new features does not enhance the performance of SVMs, revealing that such classifier can internally exploit the primary features to the maximum extent. Therefore, the classifier based on SVMs does not require an additional preprocessing step to generate more features; note that the magnitude of the failures are just linear combinations of primary features. Conversely, the substantial improvement seen in the DT classifier reveals that DTs cannot maximally exploit the information carried by the primary features and requires some pre-processing to grasp more information, which is a weakness of DTs compared to SVMs.

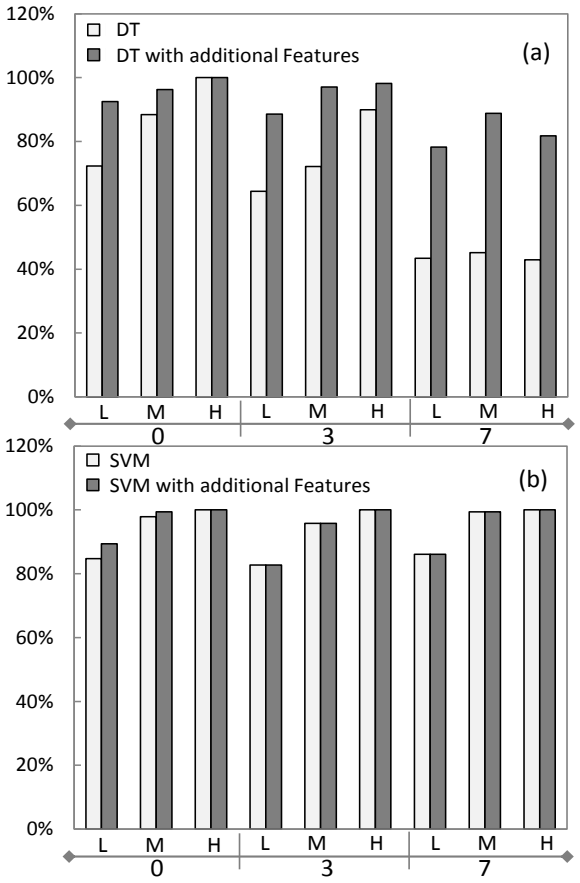


Fig. 4-14. Comparisons of classifiers with and without additional features

4.4 **Concluding Remarks**

In this chapter, we have studied the benefits of exploiting OSAs for identification and localization of filter related failures. Two different approaches for filter-related soft-failures detection and identification have been proposed and their performance based on two different ML algorithms was compared in terms of accuracy and robustness.

Multi-classifier and single-classifier, even though their performance is comparable, have notable differences in their implementation complexity. While the multi-classifier approach requires a huge dataset for the training phase, single-classifier approach requires *N* times less data, being N the maximum number of nodes an optical connection might pass-through. However, this is at the cost of pre-processing optical spectrum features, which requires the calculation of the correction masks. Table 4-2 summarizes the different of the two proposed approaches.

Table 4-2: Key characteristics and results for the considered approaches for failure detection and identification

	Pre-processing	Classification method	Training phase	Number of classifiers	Availability at the node level	Accuracy	Robustness w.r.t # of nodes
Multi Classifier	not required	SVM	requires observations of every level of filter cascading	# of nodes to support	all classifiers	good	good
Single Classifier	pre-processing of the features	SVM	requires observations of just a single level of filter cascading	1	one classifier + correction masks	good	good

In view of that the accuracy of feature-based approaches presented in this chapter when the cascading is present, the next chapter studies a different approach and proposes to circumvent the filter cascading issue in which the optical spectral themselves are preprocessed prior to interpreting them to find failures.

Chapter 5

Residual-based Spectrum Monitoring

In this chapter, we continue with the exploration of the benefits of analyzing the optical spectrum of lightpaths for soft-failure detection and identification in EONs. We present a novel approach called *residual-based* that uses the residual signal computed from subtracting the signal acquired by OSAs from an expected signal synthetically generated. Similar to the single classifier based approach presented in the previous chapter, the residual-based approach requires less data for training the classifier; instead it require two new modules to be located in the MDA agents to compute the expected signal and the corresponding residuals.

We further extend this approach to facilitate ML algorithm deployment in real network. A two-phase strategy is proposed: *Out-of-field training* is expected to use data from simulation and testbed experiments with generic equipment whereas *In-the-field* adaptation is applied to support heterogeneous equipment. Extensive numerical results are ultimately presented to compare the performance of the residual-based approach with the feature-based approaches proposed in the previous chapter in terms of accuracy and robustness. We also demonstrate the efficiency of residual adaptation mechanism for a particular use-case.

5.1 Residual-based Approach for Optical Spectrum Analysis

An alternative approach to resolve the filter cascading issue, compared to the ones in the previous chapter, is to pre-process the acquired optical spectrum by comparing it to the one that would be expected after passing the same number of filters than the signal (see Fig. 5-1). This comparison/computation produces a *residual signal* representing the differential deformation in its shape that might be due to a failure. Note that this approach does not use the FeX module proposed in

the previous chapter. In order to compute the residual signal, we consider two new modules, as shown in Fig. 5-1: *i*) the Expected Signal Calculation (ESC) and *ii*) the residual computation module.

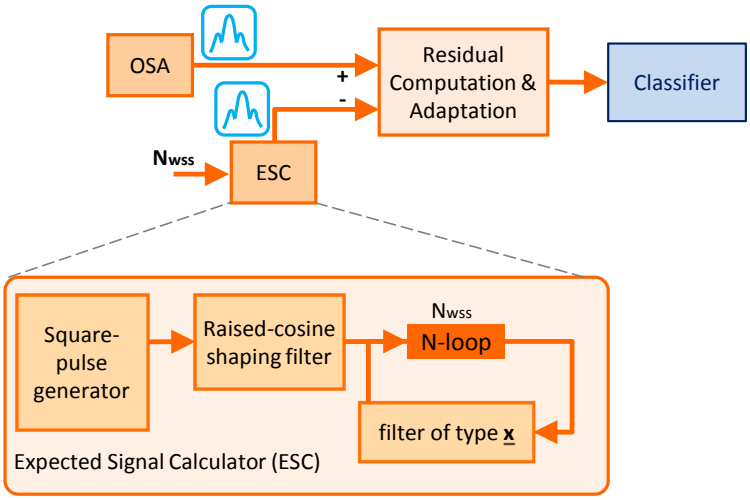


Fig. 5-1. Soft-failure detection based on residual analysis

The ESC module generates a theoretically-calculated optical spectrum emulating a properly operating lightpath. The aim of the ESC module is to synthetically reproduce an averaged noise-free version of the optical signal. In order to do so, the signal is modelled as an ideal square pulse, with $bw_{\text{-}3dB}$ equal to the baud rate of the optically modulated signal, shaped by a raised-cosine shaping filter with 0.15 roll-off factor. Then, in order to model different levels of filter cascading, a 2nd order Gaussian filter, emulating a WSS, is used (Fig. 5-1). This, results in an emulated noise-free spectrum, similar to a noise-free 100G DP-QPSK (or 200G DP-16QSPK) modulated signal (Fig. 5-2). Every time a new optical spectrum is acquired for a given lightpath, the residual computation module subtracts it from the expected signal produced by the ESC module.

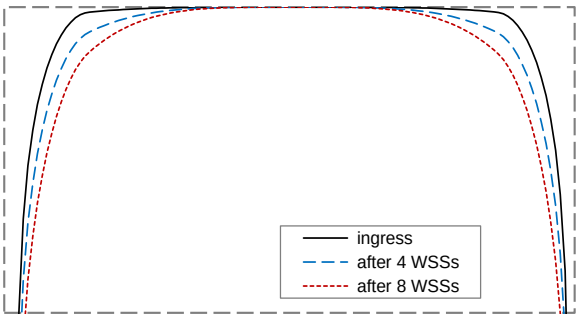


Fig. 5-2. Synthetically generated expected signals

Fig. 5-3 presents an illustrative case of the computed residual signal from the expected and the measured signals. As the residual signal experiences undesired changes at the two ends of the acquisition window of the OSA, we just consider a

central spectral window of size 34.375 GHz out of 37.5 GHz. Note that in case of a normal signal, the residual value fluctuates around a mean value along the whole signal spectrum range. To analyze the residual signal, we normalize the values, so as the mean equals 0; in such case, the most likely situation is to have as many positive as negative values. However, in the event of a filter failure, that similar proportion between positive and negative residuals will be altered.

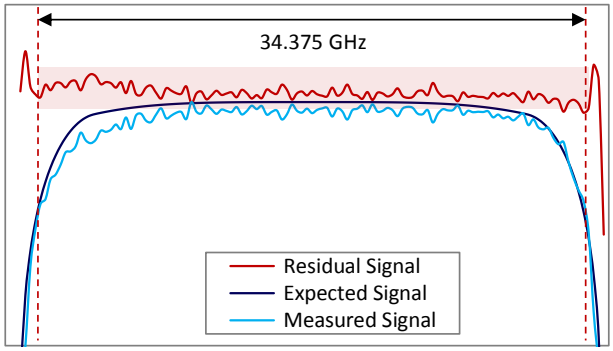


Fig. 5-3. Residual signal calculation

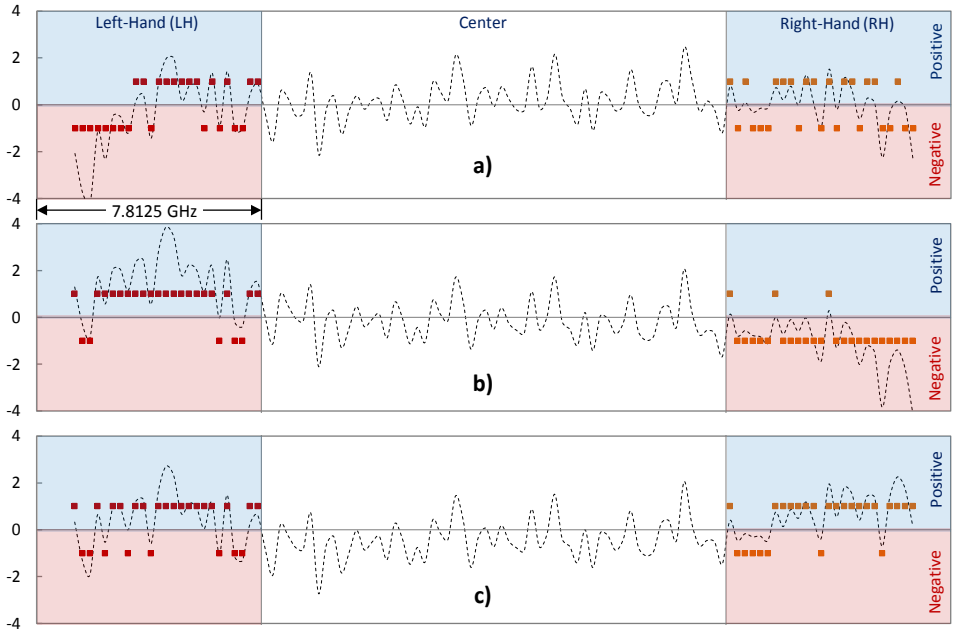


Fig. 5-4. Residual based approach workflow considering a normal signal (a), FS (b), and FT (c).

Fig. 5-4 shows an example of how residuals behave in all considered cases. Out of the whole range of the residuals, the left and the right hand-sides capture the effects of soft-failures and are the operational regions for the analysis presented next. The spectral window of the left (LH) and the right (RH) hand-sides are set to 7.8125 GHz, as they contain the sufficient number of points to capture the effects taking place in the edges. In the normal case (Fig. 5-4a), the residuals oscillate uniformly between positive and negative values (dots in Fig. 5-4 are computed from

the normalized residuals applying the *sign*() function). In the case of filter shift (Fig. 5-4b), the residuals show a clear distortion toward positive and negative parts in the LH and RH side, respectively. Finally, in the case of filter tightening (Fig. 5-4c), the residuals move toward the positive part in both the LH and RH sides.

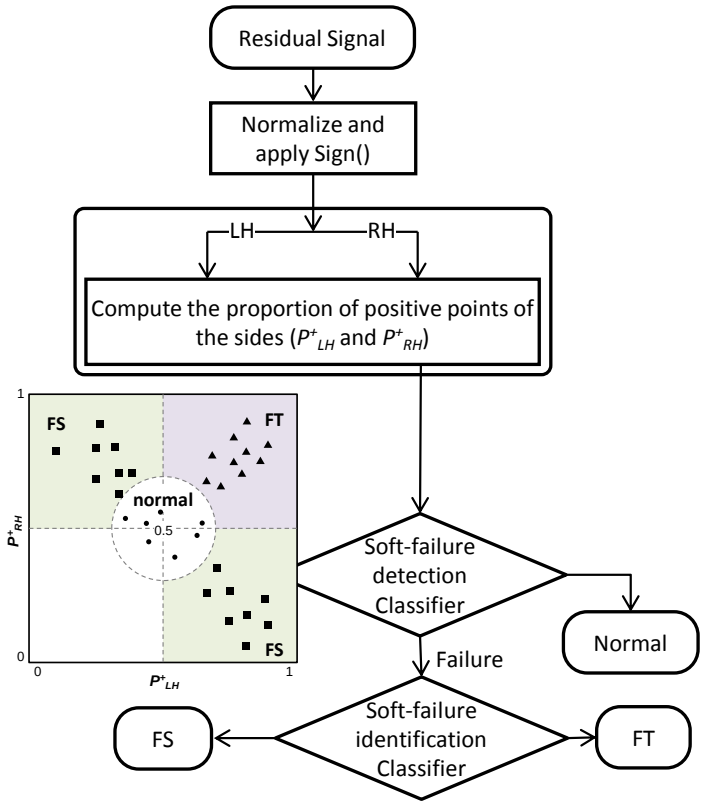


Fig. 5-5. Residual-based classification procedure

In conclusion, comparing LH and RH sides, one can predict whether the signal is normal (symmetric sides and unbiased distribution) or either it is affected by filter shift (asymmetric sides) or a filter tightening (symmetric and biased distribution).

In light of this analysis, we propose the residual-based procedure presented in Fig. 5-5, where the residuals are first normalized with respect to their mean value and centered in zero; then, the *sign*() function is applied to convert normalized residuals into points of amplitude +1 or -1. After selecting the sides, the proportion of positive points in the LH and the RH sides (P^+_{LH} and P^+_{RH}, respectively) is computed. For illustrative purposes, Fig. 5-5 plots possible observations in the semi-plane P^+_{LH},P^+_{RH}. According to the rationale previously presented, normal observations should be kept within an area centered around (0.5,0.5), FS observations should be in the quadrants (>0.5,<0.5) or (<0.5,>0.5), while FT observations should be in the quadrant (>0.5,>0.5), both FS and FT outside the normal area. Therefore, the coordinates of the observations in the semi-plane P^+_{LH},P^+_{RH} can be used as features for DT and SVM -based classifiers. In the

proposed procedure, two classifiers are trained; the first one for soft-failure detection and the second one for its identification.

5.2 Facilitating ML Algorithm Deployment Using Residual Signals

The premises described in the previous section consider one single filter type, which limits the deployment of ML approaches to real operator networks that usually consist of equipment from different vendors. The most straightforward solution to overcome this limitation is to have different models being trained upon various types of filters that might be available in the network. Nonetheless, it makes the training phase very complex and data-hungry. Yet, it will not be easy to comprehend the sequence of filters a priori and the responses of a slightly non-identical filter in the network might not be very well detected, necessitating even more combination of models to have an appropriate generic model.

Here, we propose to extend the premises described in the previous section to facilitate ML algorithm deployment in real networks; it consists in:

- training one single accurate and robust ML model based on simulations and/or experiments carried out in laboratory or testbed facilities and
- devising a proper adaptation mechanism that makes adjustments on the data for the specific signal being analyzed, which might have traversed different filter types along its route from the transmitter.

Note that this strategy also facilitates the introduction of new filter types, as current vendors deploy new equipment releases in the network. The *residual-based* approach, due to its dependency on the synthetic behavior of the filter responses, has potential characteristics to get adapted to different types of filter.

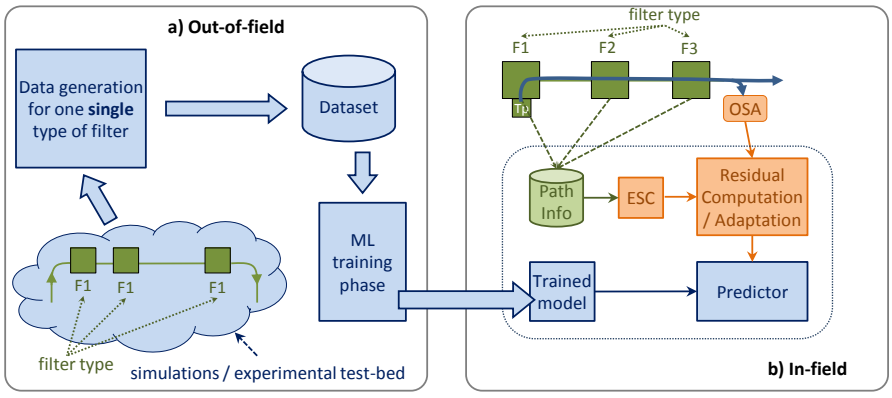


Fig. 5-6. Out-of-field ML training and in-field model adaptation

Fig. 5-6 illustrates the proposed out-of-field ML training and in-field model adaptation strategy. Scenarios with one single type of filters (labeled F1 in Fig.

5-6a) are considered in simulation and/or lab experiments to produce a large dataset that is used for ML training purposes. When the ML model is deployed in the field, an adaptation procedure takes into account the specific types of filter that a given signal has passed through (Fig. 5-6b). The in-field adaptation is performed in *i*) the ESC module by considering the specific filters that the signal has passed through; see three filter transfer functions in Fig. 5-2, and *ii*) the residual computation module that normalizes and adapts the residuals for the signal under analysis.

In the procedure described in Section 5.1, the calculated residual is normalized with respect to the mean value of the central part of the residual, so the mean becomes 0. This normalization approach is operational when the same type of filter exits in both out-of-field training and in-the-field operation of the ML algorithm. However, when applied to other filter types it does not work well. In view of this, we propose an adaptation procedure (Fig. 5-7) that consists in dividing the residual signal in three segments (Fig. 5-8) and apply different normalization methods to every segment, reflecting the filter characteristics; the normalization reference of every segment is obtained by applying linear regression to the un-normalized version of the residual signal obtained for that segment.

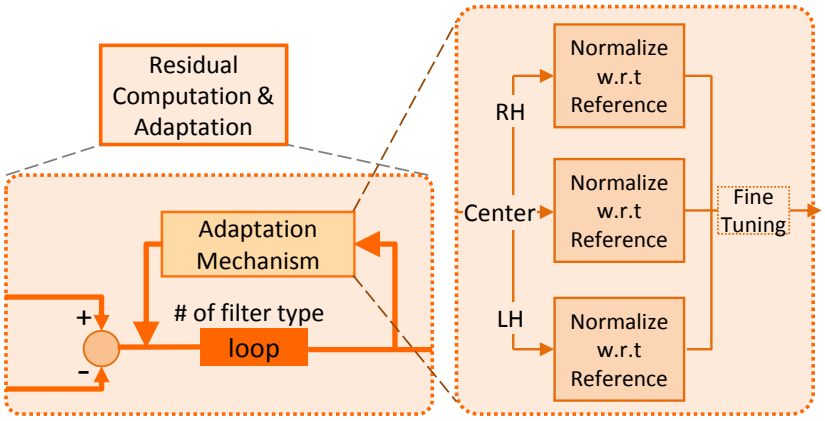


Fig. 5-7. Residual computation and adaptation mechanism

When the signal arrives to the residual computation and adaptation module, the module also receives three different normalization references considering the filter characteristics in the form of polynomials of order 1. For this stage, the number of adaptation mechanism loop equals to the number of filter types that the lightpath has passed through. By subtracting every segment of the un-normalized residuals from the corresponding normalization reference, a filter-agnostic residual signal is obtained.

Note that, as the amount of filter cascading effect depends on the transfer function of the filter, there might be an undesirable power shift in the residual signals when the lightpath traverses different filter types. This power shift is compensated in the fine tuning step. The amount of power shift can be computed locally assuming that

the mean value of the residual remains zero when the signal is in proper operation mode. Ultimately, a single classifier trained with the measurements collected in the lab based on a reference filter type, can be used for optical spectra experiencing filtering effects from different types of filter.

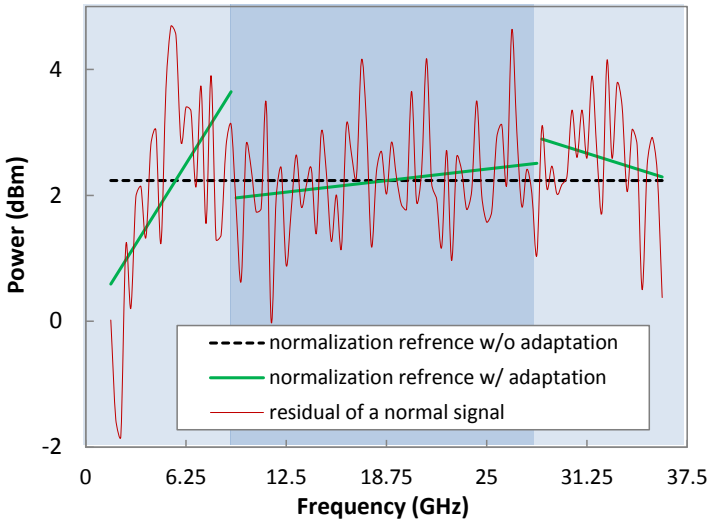


Fig. 5-8. Normalization references for the 4th order Gaussian

The result of this adaptation method is illustrated in Fig. 5-9. The residual signals of a lightpath passing through three different types of filters with Gaussian transfer function of order 2, 3, and 4 are illustrated in Fig. 5-9a. Normalization shifts the residuals so its mean to be 0 (Fig. 5-9b). Note that the differences among residuals are clearly seen at the edges, whereas they are virtually identical in the central part before and after normalization. Adaptation focus on compensating the effects of the different filters and the results are clearly visible at the edges (Fig. 5-9c); note that the most relevant parts of the residuals to detect filter-related soft-failures are that of the edges. As shown, even though the signals pass through different types of filters, they result in an identical residual signal, removing the filter-dependent characteristics of the residual signal.

5.3 Illustrative Results

In this section, we first numerically compare the performance of the residual-based approach described in Section 5.1 with the feature-based approaches presented in Chapter 4. The same transmission set-up described in Section 4.3.1 is used here to generate the optical spectrum database required for training and testing the proposed algorithms.

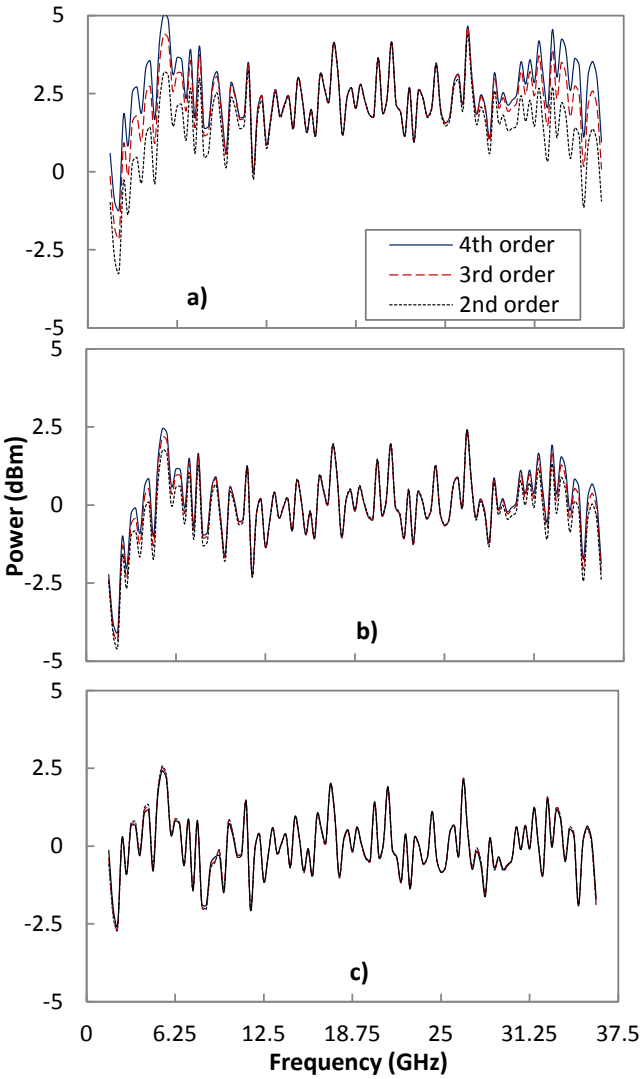


Fig. 5-9. Un-normalized residual (a), normalized w/o adaptation (b), and normalized with adaptation (c)

5.3.1 Comparison of Residual-based and Feature-based Approaches

Let us first focus on the residual-based approach and analyze the potential of using P^{+}_{LH} and P^{+}_{RH} as discriminatory features for training the classifiers. Fig. 5-10 plots the samples of the dataset used for training the residual-based classifiers at N1. It is clear, in view of Fig. 5-10, that the selected features allow classifying the different cases easily. Although not shown in detail, we first compared DT and SVMs classifiers and concluded that SVMs perform better also for this approach. Therefore, the results presented next compare the performance of all three approaches using SVM-based classifiers.

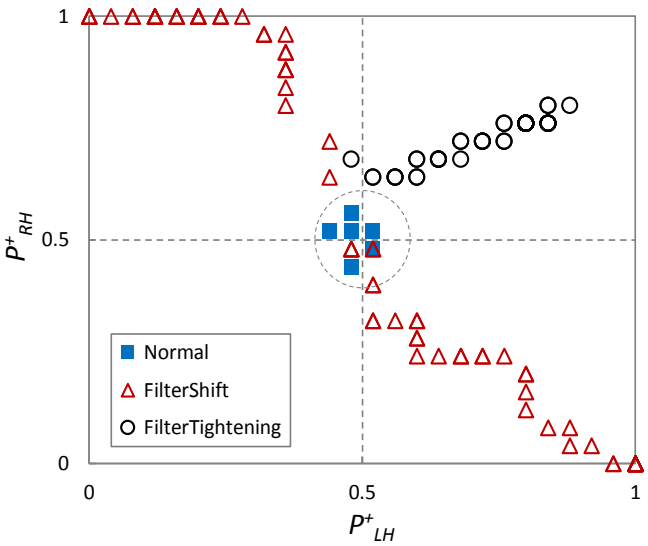


Fig. 5-10. P^{+}_{LH} and P^{+}_{RH} features for failure detection and identification

Fig. 5-11a-b show the accuracy of detecting FS and FT solely at node N1. As illustrated, multi-classifier and single-classifier approaches show identical performance as they are essentially the same at node N1, where no filter mask is necessary to be used. However, the residual-based approach outperforms feature-based approaches for the detection of FS and FT. The performance of the residual-based approach is noticeable for detecting soft-failures, as it reaches 100% accuracy for 3 GHz FS magnitude and 5 GHz FT magnitude, 2 GHz smaller than the other two approaches.

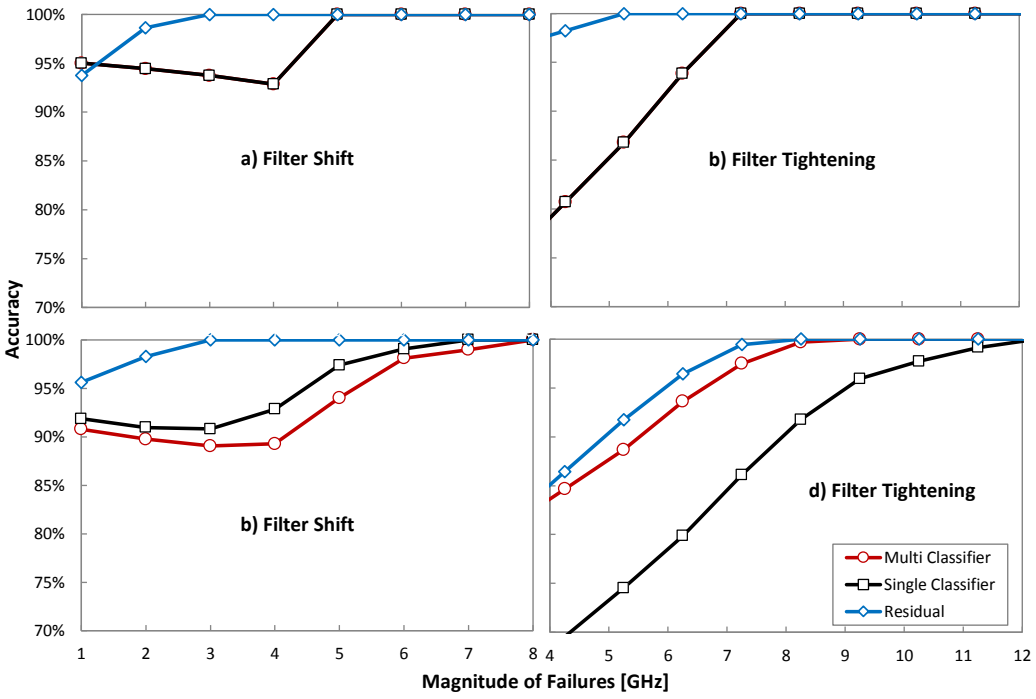


Fig. 5-11. Accuracy of the different approaches

Let us now look at the average accuracy of the approaches over all the nodes; the results are shown in Fig. 5-11c-d for FS and FT failures, respectively. The residual-based approach remains the best solution by far for detecting and identifying both, FS and FT failures. In contrast, the multi-classifier-based approach shows the worst overall accuracy for FS failures, as the features selected for classification get very close to each other and it becomes very difficult for the support vectors to distinguish them perfectly. Interestingly, the situation is different for FT failures, where the multi-classifier approach performs better than the single-classifier.

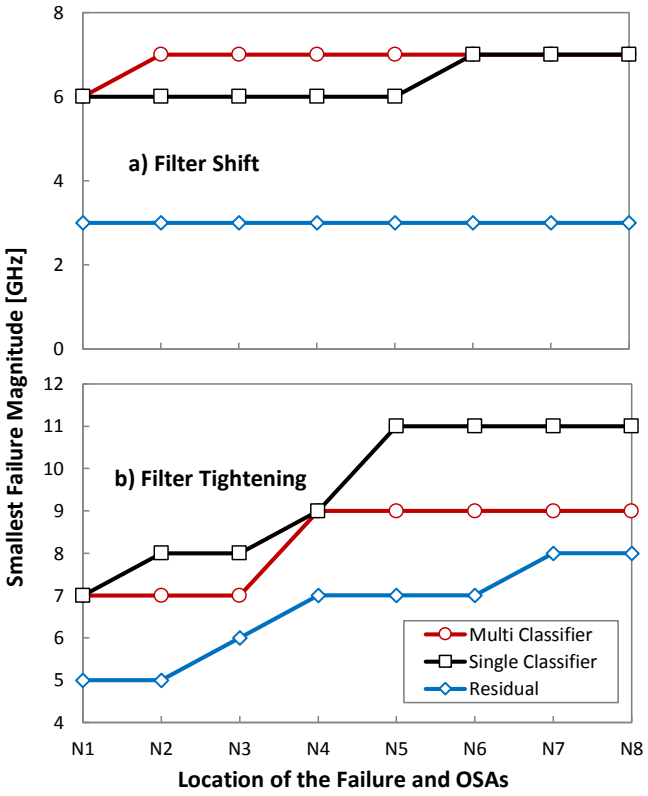


Fig. 5-12. Robustness of the different approaches

In addition to classification accuracy, the *robustness* of the approaches with respect to the number of traversed nodes is of paramount importance for practical implementations. Hence, let us compare the robustness of different approaches in terms of the smallest failure magnitude after which the classification accuracy reaches 100%. Plots in Fig. 5-12a-b represent such robustness in terms of the location of FS and FT failures, respectively. Note that, points for N1 in Fig. 5-12a-b correspond to the first points in Fig. 5-11a-b where 100% accuracy is achieved. As observed, the residual-based approach shows the highest level of robustness compared to the feature-based ones for both FS and FT failures. It can be understood that the residual-based approach is robust regardless of the location of the failure as it perfectly detects and identifies failures with magnitude above the values in Fig. 5-12.

5.3.2 The Efficiency of Residual Adaptation Mechanism

Let us now to discuss numerical results to demonstrate how the proposed adaptation mechanism enables the residual-based approach to be applied to optical spectrum of a signal after passing through different types of filters in the network. For the experiments, we used the same VPI Photonics setup used before. Aiming at emulating failure scenarios, we modified the characteristics of the 2nd WSS of every node in the setup; its bandwidth and central frequency were modified to model FT and FS failures, respectively. A large dataset of failures was collected by inducing failures of magnitude in the range [1-8] GHz for FS and in the range [1-15] GHz for FT. We configured optical filters to be 2nd order Gaussian for training and re-configured them to become 3rd and 4th order Gaussian for testing, where the same failure scenarios were simulated.

We look firstly at the benefits of applying the adaptation mechanism for identifying the normal cases. We found that accuracy (number of correctly detected cases over the total number of cases) is poor when no adaptation is applied and becomes perfect with residual adaptation. Next, we looked at the benefits of applying residual adaptation for detecting failures. Three cases were studied:

- *2nd order* for both out-of-field training and in-field testing; note that no adaptation is needed (the case in the previous section),
- *3rd order*, in which 2nd order filters were used for training and 3rd order filters w/ adaptation were used for testing.
- *4th order*, in which 2nd order filters were used for training and 4th order filters w/ adaptation were used for testing.

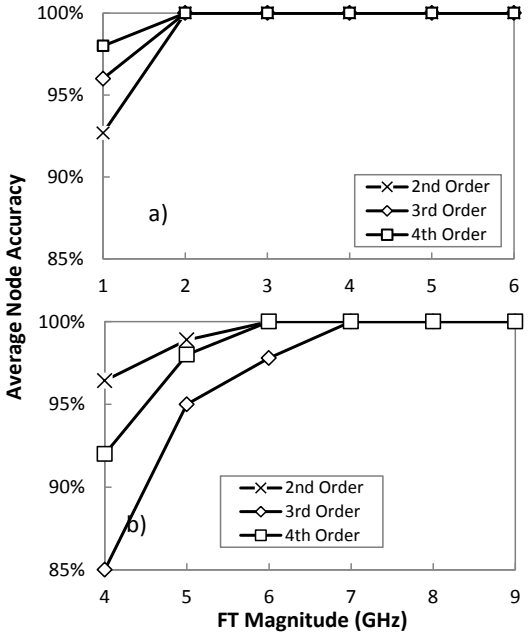


Fig. 5-13. Average node accuracy w.r.t failure magnitude for a) FS and b) FT.

The results are reported in Fig. 5-13, where Fig. 5-13a and Fig. 5-13b show the average node accuracy of identifying FS and FT, respectively, for failures in all 8 nodes and varying levels of failure magnitudes. The accuracy is promising for all the cases under study, even though it degrades for very small magnitudes in which the spectrum looks like normal cases; in fact, failure detection is 100% in all cases being the failure identification step the cause of the reduced accuracy (Table 5-1).

Table 5-1: Results Comparison

Scenario	Failure Detection	Failure Type Identification	
		Min FS Magnitude	Min FT Magnitude
only 2nd or 4th order	100 %	2 GHz	6 GHz
only 3rd order	100 %	2 GHz	7 GHz
mix of 2nd and 4th order	100 %	5 GHz	7 GHz

To highlight the impact of cascaded nodes, Fig. 5-14 presents the average accuracy for FS and FT with respect to the node where the failure occurs; failure magnitudes of range [1-4] GHz for FS and [4-7] GHz for FT were considered. As shown, the accuracy drops at the very last nodes as a result of accumulated filter cascading effects makes very challenging to distinguish between different cases.

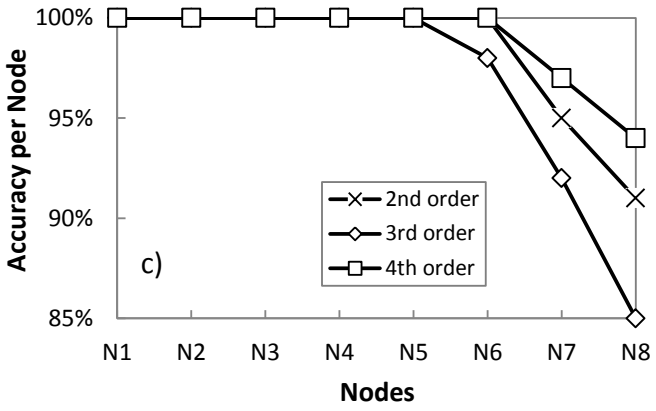


Fig. 5-14. Accuracy per node w.r.t the sequence of cascaded nodes (c).

Ultimately, the efficiency of the algorithm for transmission system with two different filter types was evaluated. To this end, we modified the above described setup to have 2nd order Gaussian filters in the first 4 nodes and 4th order Gaussian filters in the last 4 ones. As reported in Table 5-1, failure detection accuracy is 100% while performance degradation happens when failure identification is executed. As a result, the minimum failure magnitude to be detected with 100% accuracy is 5 and 7 GHz, for FS and FT, respectively, just a bit higher than in the case of one single filter type, which validates the performance of the proposed residual adaptation method.

5.4 Concluding Remarks

In this chapter, we have continued exploiting the benefits of OSAs for identification and localization of filter related failures. A new approach for filter-related soft-failures detection and identification, called residual-based, has been proposed and its performance was compared in terms of accuracy and robustness against the two approaches proposed in the previous chapter.

The residual-based approach, which is based on a single classifier strategy, significantly outperforms the feature-based ones and brings down the complexity of training phase compared to multi-classifier approach. However, it requires two additional modules to be available in the MDA agent. Table 5-2 summarizes the key differences of the three approaches.

To conclude, it is beneficial to bring data analytics as close as possible to the MDA agents as the complexity of the proposed modules is low enough to be integrated in programmable units that are expected to be available in the future whitebox-based technologies for switching and Tp optical nodes.

Table 5-2: Key characteristics and results for the considered approaches for failure detection and identification

	Pre-processing	Classification method	Training phase	Number of classifiers	Availability at the node level	Accuracy	Robustness w.r.t # of nodes
Multi Classifier	not required	SVM	requires observations of every level of filter cascading	# of nodes to support	all classifiers	good	good
Single Classifier	pre-processing of the features	SVM	requires observations of just a single level of filter cascading	1	one classifier + correction masks	good	good
Residual Based	pre-processing of the optical spectrum	SVM	requires observations of just a single level of filter cascading	1	one classifier + ESC and residual signal computations modules	very good	very good

We have also shown that ML algorithm can be trained out-of-field with measurements from testbeds and/or simulations using one single reference filtering solution (possibly belonging to a single vendor). A reliable in-field adaptation mechanism is demonstrated to enable heterogeneous filtering solutions (belonging to different vendors).

While filter related failures may happen in SSON, they are not of any concern in FON. However, other different failures might happen in FONs that will required the adaptation of the approaches proposed in this and the previous chapters to the characteristics of FONs.

Chapter 6

Monitoring of Filter-less Optical Networks

The previous chapter targeted the use of optical spectrum monitoring to detect and identify filter-related problems in SSON. In this chapter, we are going to present a a real-time surveillance system in FON, which exploits data analytics and cost-effective OSAs. Specifically, we present a set of novel signal identification, classification and tracking algorithms for FONs and study the most proper OSA resolutions for their efficient deployment in the networks. In this regard, we propose to use one OSA in FONs to continuously scan the whole C-band and utilize the acquired spectrum to perform signal identification and classification tasks.

Moreover, we propose two different optical signal tracking solutions, which are launched upon a notification from the node controller;

- *feature-based tracking*, which uses features extracted from the optical spectrum, and
- *residual-based tracking*, which uses a set of residual signals calculated by subtracting the measured signal by the OSA from an expected version of the signal.

We propose two different flavors of the feature-based approach; one of them uses the features directly extracted from the optical signal, while the other one manipulates the features to generate super features resolving some of the issues of the former approach. In the case of the residual-based approach, in addition to its excellent *individual* signal tracking, it is enriched with the capability of using tracking information from one signal to enhance the tracking accuracy of neighboring signals (we call this *contextual*).

In order to demonstrate the performance of our proposals, we challenge them in two different scenarios of a FON:

- we simulate a direct detection PAM4 transmission system and apply the proposed algorithms to the PAM4 optical spectra, and
- we apply the algorithms to a set of experimental measurements of a QPSK transmission system.

6.1 Motivation of Optical Monitoring in FONs

In contrast to filtered optical networks, FONs eliminate or minimize the number of active switching elements in the optical line systems. An example of FON connecting packet nodes is presented in Fig. 6-1. In FONs, traffic is broadcasted throughout the network; for instance, in Fig. 6-1 four lightpaths are created: R1->R4 (labeled 1), R2->R3 (2), R2->R5 (3), and R3->R5 (4).

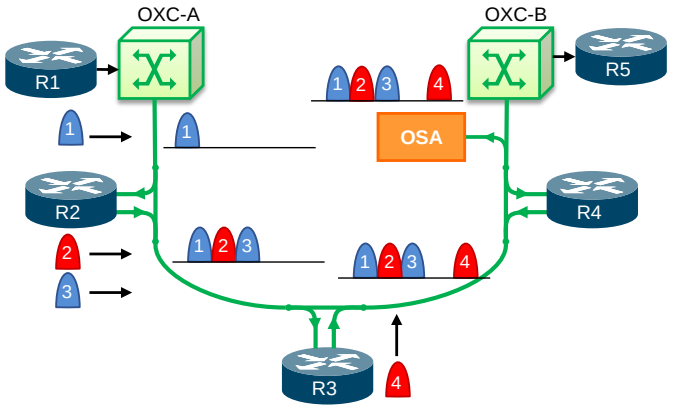


Fig. 6-1. Example of a FON

Let us imagine that Tps for lightpaths 2 in R2 and 4 in R3 experience a Laser Drift (LD) failure shifting the signals to the right in the spectrum. In this case, signal 3 might be affected as the spectrum of signal 2 overlaps it, whereas signals 2 and 4 may be affected if the receiver is not capable of tracking their Central Frequency (CF) and misalignment between signal and the optical filter at the receiver occurs, as introduced in Chapter 1.

In contrast to coherent transmission systems, in PAM4, a small detuning of the transmitter leads to a noticeable performance penalty in the receiver. DD receivers are not capable of tracking the laser wavelength of the transmitter and any amount of LD results in misalignment between the signal launched by the transmitter and the optical filter in the receiver (see Fig. 6-2a). LD can potentially produce signal overlap when channel spacing between two neighboring signals is reduced (see Fig. 6-2b). Misalignments and overlaps ultimately introduce significant performance penalty.

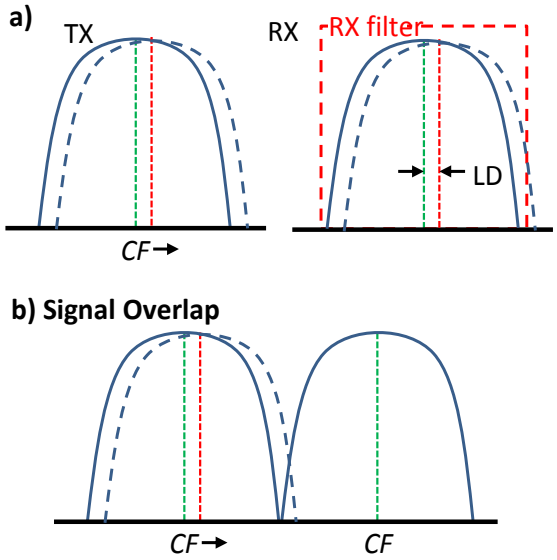


Fig. 6-2. Signal misalignment and overlap

In order to resolve these issues, one conservative solution could be assigning larger channel spacing thus, deteriorating even more the spectral efficiency of FON. A more effective solution is to consider a reliable and robust spectrum surveillance and optical signal tracking procedure to detect LDs accurately, so Tps can be adequately retuned. Consequently, margins (e.g., channel spacing) can be reduced while assuring proper Tp operation. Note that Fig. 6-8 and Fig. 6-12 represent this alternative lecture; we could reduce channel spacing provided that very stable systems are deployed in the network.

The proposed spectrum monitoring approach has an advantage while exploited for FONs, which is the small number of OSAs required for real-time network spectrum monitoring; note that in a similar SSON, one OSA per link (five in total) would be required. In this chapter, we propose to use one single OSA installed in the last span, where all signals in the FON can be acquired. Captured spectrum needs to be analyzed real-time so active lightpaths in the FON are monitored and prompt actions are taken before a properly operating lightpath becomes affected by a failed Tp.

Note that the frequency range of a signal might not be exactly determined and slightly change along lightpaths' lifetime. Therefore, an algorithm examining the captured optical spectrum cannot select a frequency range in the whole C-Band acquired by an OSA and focus on analyzing it in the hope that the whole spectrum of a target lightpath and only of such lightpath is confined within that frequency range.

In consequence, we propose algorithms that periodically scan the whole C-band and rely on an ordered list of lightpaths, including relaxed frequency ranges for each one, obtained from the SDN controller; the scan process is intended to ensure that

signals in the network and lightpaths in the list match in terms of frequency ranges. Any found difference (i.e., signals not in the list and lightpaths not in the FON), as well as detected anomalous signal CF shifts that might end in impacting neighboring lightpaths are reported to the SDN controller. Analyzing the current signals' spectrum allocation and lightpaths information from the controller, we thus aim at checking whether each signal is confined within the frequency range allocated to a lightpath (*normal* signals); conversely, three anomalies can be identified (illustrated in Fig. 6-3a), namely:

- a signal is partially out of the spectrum allocated to a lightpath (*outOfRange*);
- a signal is in a spectrum range no allocated to any lightpath (*unknown*); and
- no signal has been detected in the spectrum allocated to a lightpath (*missing*).

The detection of any of these anomalies triggers a notification with *critical* severity level to the controller, whereas *normal* signals need to be tracked afterwards to predict a potential anomaly.

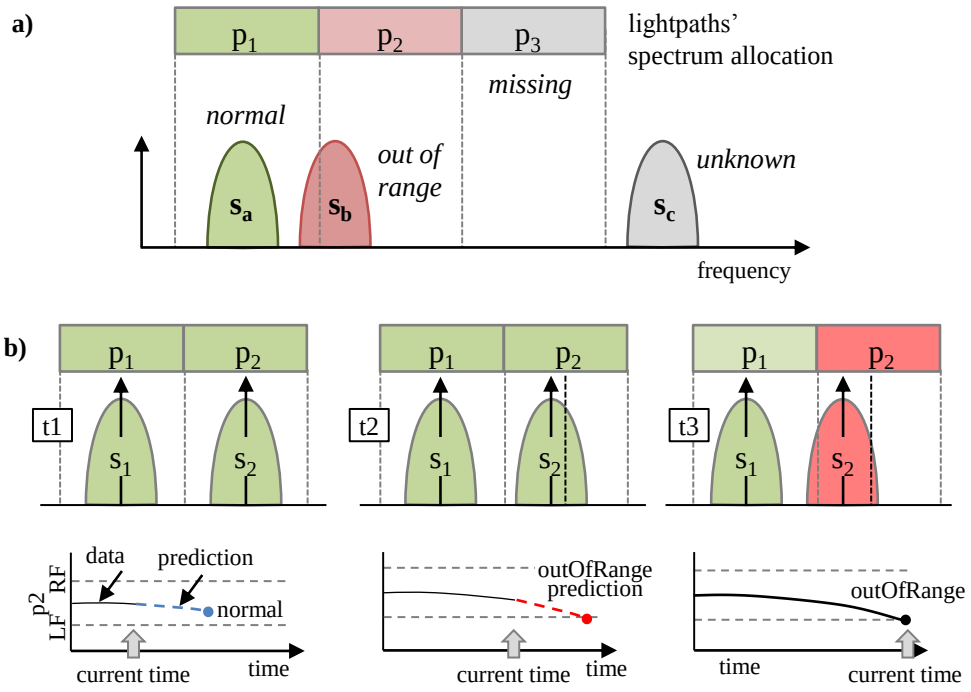


Fig. 6-3. Signal classification example (a) and anomaly detection (b)

6.2 Signal Identification and Classification

The proposed procedure starts when a new C-band scan is acquired by the OSA. The first step is to detect the allocated spectrum to each signal; by using the derivative of the power w.r.t. the frequency, as we introduced in Chapter 4. The sharp power rising at the left frequency edge followed by the power falling at the right frequency edge of each signal in the spectrum can be detected. Next, the algorithm in Table 6-1 is used to classify the set of identified signals S w.r.t to the list of lightpaths P.

Table 6-1: Signal Classification Algorithm

INPUT *S, P*

OUTPUT *normal*, *outOfRange*, *missing*, *unknown*

```
1:  normal=outOfRange=missing=unknown←Ø
2:  ID←getAllIds(P)
3:  S'←Ø
4:  for each s∈S do
5:    P'←findOverlappingLightpaths(P,s)
6:    if P'≠ Ø then unknown←unknown ∪ {s}
7:    else if |P'|==1 and totalOverlap(s,P') then
8:      normal←normal ∪ {<P'.getId(), s>}
9:      ID←ID \ {P'.getId()}
10:   else s.I← P'.getId()
11:     S'← S' ∪ {s}
12: if S'≠ Ø then
13:   for each s∈S do
14:     I←{s.I} ∩ ID
15:     outOfRange ← outOfRange ∪ {<I.first,s>}
16:     ID←ID \ I.first
17: if ID≠ Ø then missing←missing ∪ ID
18: return normal, outOfRange, missing, unknown
```

After some initializations (lines 1-3 in Table 6-1), the algorithm iterates on the signals to find the lightpaths where the allocated spectrum includes part of their range (lines 4-5); if no lightpath is found, the signal is classified as *unknown* (line 6), whereas it is classified as normal if the allocated spectrum of just one lightpath totally overlaps the signal (lines 7-9). Otherwise, signals are classified as *outOfRange* and assigned to the first overlapping lightpath if more than one exists (lines 10-16). Finally, the set of *missing* lightpaths (if any) are obtained and the classification results eventually returned (lines 17-18).

Non- *normal* signals trigger notifications to the controller and they can be discarded for further analysis. The next step focuses on tracking *normal* signals to predict any possible violation of their spectrum allocation that could impact on neighboring signals. The proposed optical signal tracking approaches are detailed in the next section.

6.3 **Optical Signal Tracking**

In this section, we present two different approaches to track the movement of an optical spectrum.

6.3.1 Feature-based Tracking

This approach relies on extracting some meaningful features describing the key characteristics of the optical spectrum. In this approach, we consider the FeX module introduced in Section 4.2.2.

Considering the value of the features obtained by FeX module, we can follow two different approaches to track the movement of an optical signal, which is essentially equivalent to the amount of LD experienced by that signal. The first one (named as *Individual Feature*) relies on using a few meaningful features calculated directly from FeX module, while the second one (named as *Super Feature*) relies on modeling a LD estimator using a multiple linear regression model. The *Individual Feature* relies on a single identical feature of the optical signal itself, while the *Super Features* does not rely solely on the extracted features, but uses a multiple linear regression model to transform the extracted features from FeX to a more informative one, which turns out to be a more accurate estimation of LD using a combination of several feature. Both approaches are detailed below.

6.3.1.1 Individual Feature

In this approach, we use the relevant features extracted by FeX module to estimate the LD and track the evolution of the signals. LD is defined as the difference between the expected *CF* of the signal and the estimated one ($LD = CF_{exp} - CF_{est}$). Using the outputs of FeX, *CF* can be calculated by subtracting/adding half of the $bw()$ from/to the cut-off points of different power levels at the right/left side; for instance CF_{-3dB} can calculated in one of these ways $CF_{-3dB} = f_{2(-3dB)} - bw_{(-3dB)}/2$ or $CF_{-3dB} = f_{1(-3dB)} + bw_{(-3dB)}/2$. While CF_{est} is useful for estimating the LD, power levels at the left and right frequencies (i.e. P_{f1} and P_{f2}) are more relevant to detect signal overlaps. Therefore, these relevant points can be used to track the evolution of the signal with time and eventually to predict whether it is likely to exceed the spectrum allocation within a given future time window.

An example of this procedure is illustrated in Fig. 6-3b, where signal *s2* is gradually approaching neighboring signal *s1*. In this case, the prediction of *s2* P_{fl} at time *t2* states that it will exceed its spectrum allocation and thus, a notification with *warning* severity level is triggered towards the controller before an *outOfRange* anomaly is detected (which actually happens at time *t3*). This approach has low complexity as it just requires the output of FeX module and does not require modeling any estimator a priori. However, it has a drawback, which is the dependency of its results on the resolution of the OSAs.

6.3.1.2 Super Features

While in the previous approach the LD can be estimated by performing simple addition/subtraction of the outputs of FeX module, in this approach, we need to manipulate the outputs of the FeX module to obtain an accurate model for estimating the amount of LD. This approach requires a small database of optical spectra including various levels of LDs and the corresponding relative points of each spectrum returned by FeX module to model a multiple linear regression that can be later used to estimate the LD. Even though this approach becomes a bit more complex as compared to the previous one, due to the need for a database and large number of required inputs from FeX, it performs much better and resolves the OSA resolution dependency of the previous approach.

6.3.2 Residual-based Tracking

In the residual based tracking, we face the problem from another perspective. Our proposal relies on computing an expected version of the signal (as detailed in Chapter 5) and uses that as baseline to track the movement of the signal. Therefore, the LD estimation module requires the expected shape and the location of every signal in the spectrum. As expected signals need to be generated considering the characteristics of the lightpath under study, for the sake of clarity; we describe the procedure using an example of a PAM4 optical signal. The measured spectra of PAM4 signals and their corresponding expected ones are shown in Fig. 6-4, where a 1.2GHz-resolution OSA was emulated by averaging power values on windows of such width. Note that the optical carrier vanished so it cannot be used for LD estimation.

Fig. 6-4a shows a properly configured PAM4 signal, whereas Fig. 6-4b shows the case when the signal experiences 600 MHz of LD. The next step is to subtract the expected signal from the measured one, which produces a third signal called *residual* signal. If the measured and expected signal are on the same CF, the residual signal oscillates around 0 (Fig. 6-4a). On the contrary, a positive or negative slope is observed in the residual signal, depending on the direction of the drift (Fig. 6-4b). For illustrative purposes, Fig. 6-5a plots the residual signals for LD from 0 to +4.8 GHz with 600 MHz step size.

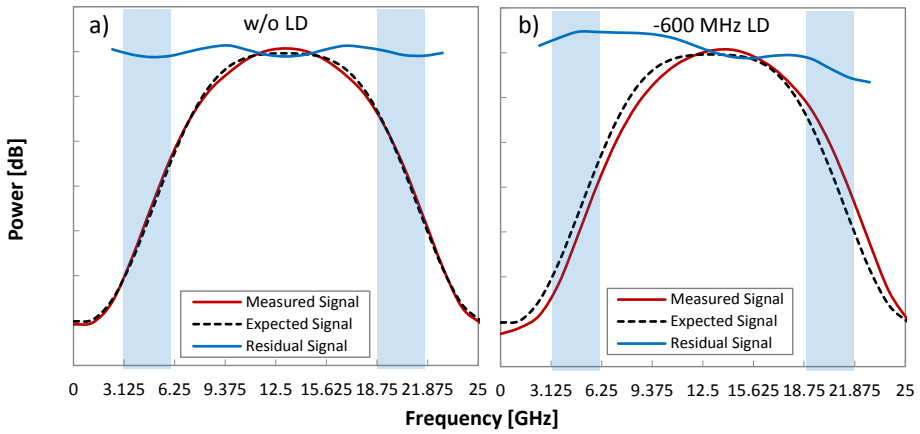


Fig. 6-4. PAM4 optical spectrum a) w/o LD, b) w/ 600 MHz LD

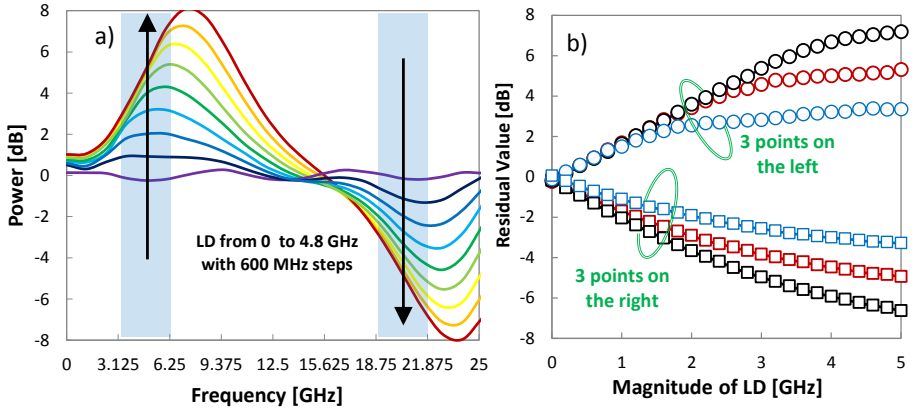


Fig. 6-5. Residuals for different LD magnitudes (a) and the evolution of residual features wr.t. the LD (b).

For modeling the LD estimator, the 3.125 GHz portions of left and right edges highlighted in blue in Fig. 6-5 were considered as features. These portions are essentially the edges of the spectrum that are expected to be affected due to the LD. Note that in the particular case of Fig. 6-5, where the traces are based on OSAs of 1.2 GHz resolution, these portions contain only 3 different power-frequency points. The evolution of these points w.r.t the magnitude of LD is illustrated in Fig. 6-5b. The final step is to consider these features and apply multiple linear regression to obtain a model for LD estimation.

When the maximum expected LD is small compared to the spectral occupation of the signal (in Fig. 6-4, we consider the maximum 5 GHz of LD compared to 25 GHz spectral occupation of PAM4 small), a single multiple linear regression model is sufficient to perform the LD estimation task accurately. However, if the expected shift in the signal becomes comparable with the spectral occupation of the channel (as in the case shown in Fig. 6-6), more than one model is required to have an accurate estimation of the shift. Considering such large shift makes sense, for

instance, when the target task is channel spacing reduction. Fig. 6-6 shows optical spectrum of 100G QPSK-modulated optical spectrum as well as its corresponding expected version and the calculated residuals.

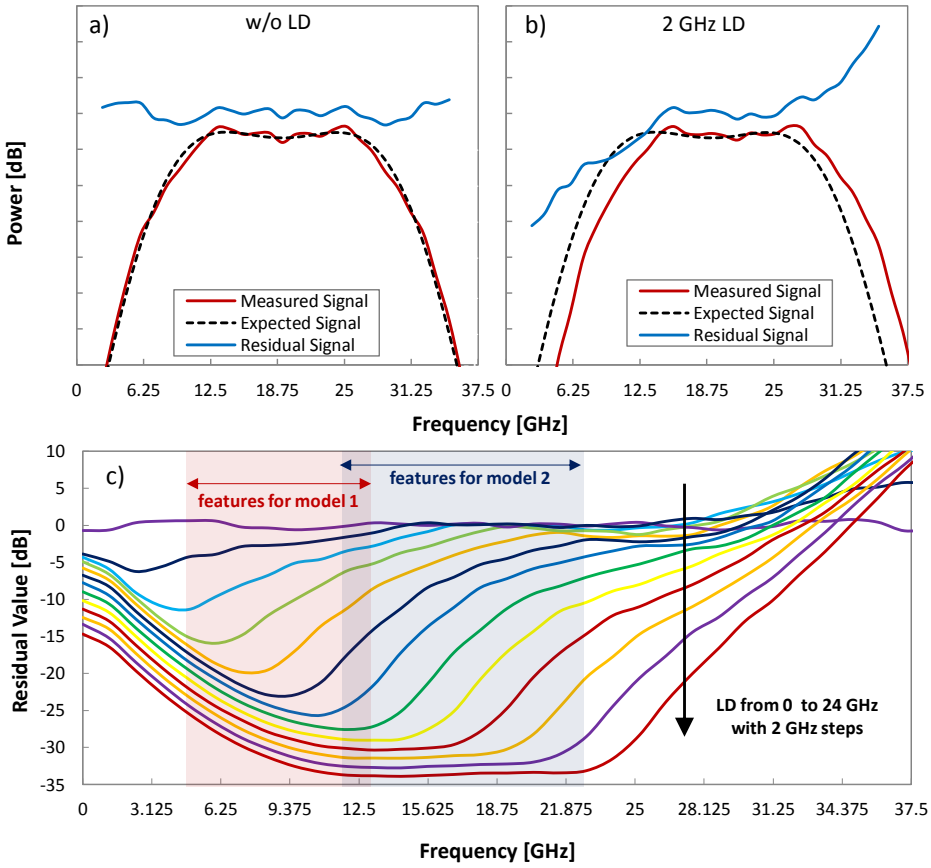


Fig. 6-6. QPSK optical spectrum captured by a 1.2 GHz OSA a) w/o LD, b) w/ 2 GHz LD, and c) residuals for different LD magnitudes.

Fig. 6-6 a shows the signal when LD is assumed to be zero, while Fig. 6-6b shows the case when 2 GHz of LD is considered. In Fig. 6-6c, the residual traces are illustrated for LD magnitude of 0 to 24 GHz with step size of 2 GHz. As shown, the residual signal at lower frequencies seems to be saturated and does not contain useful information for the LD estimation of high magnitude. Therefore, in such scenarios, different multiple linear regression models can be obtained and called depending on the situation. In the case of Fig. 6-6, two models are enough; one for LD magnitude below 10 GHz and one for magnitudes higher than that and less than 25 GHz. The spectral regions considered as features for both models are highlighted and the evolution of features of both portions are plotted in Fig. 6-7.

The above procedure works well when the signals are *individually* considered. However, in the case of signal overlap like between signals 2 and 3 in Fig. 6-1, that procedure needs to be improved, as it finds LD in signal 3, which is actually properly configured. To avoid the effects of spectrum overlap, the LD estimation

procedure analyzes the spectrum forward from left to right and considers *contextual* information; once it detects LD in one signal (e.g., in signal 2), it analyzes the following one (e.g., signal 3) considering the actual position of the previous one.

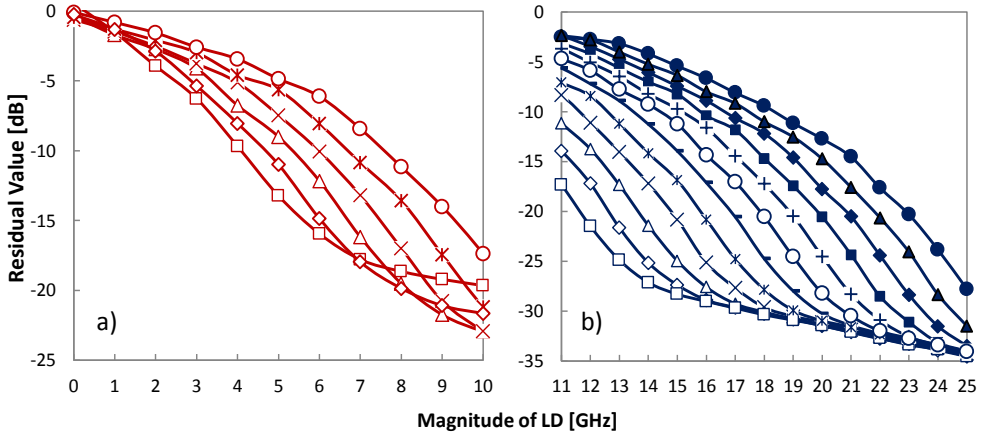


Fig. 6-7. The evolution of residual features w.r.t the magnitude of LD

This approach is a bit more complex than the previous ones, as it requires an additional module to calculate the expected signal for every lightpath. Additionally, like the *Super Features* based approach, it requires a database of optical spectra for obtaining the multiple linear regression model.

6.4 **Illustrative Results**

In this section, we present results demonstrating the efficiency of the proposed algorithm. We apply the developed algorithms to two different set of observations:

- obtained by performing a set of simulations for a PAM4 transmission system,
- obtained by setting up a lab experiment for a QPSK transmission systems.

Note that, both of the above approaches were followed to generate enough data to model and then evaluate the proposed approaches.

6.4.1 PAM4 Scenario

In the first scenario, we simulate a PAM4 system. For our experiments, we considered a 12.5 GBaud (25 Gb/s) PAM4 for cost-effective metro networks, where the electrical shaping filter of the transmitter is modeled as a 12.5GHz Root-Raised-Cosine (RRC) shaping with mild roll-off factor of 1, driving a Mach-Zehnder Intensity Modulator with 30dB extinction ratio. Propagation along 50km of uncompensated standard G.652 SMF is considered, including EDFAs to

compensate for losses; OSNR is set equal to 28dB. At the receiver, the desired channel is extracted through a 25GHz optical filter modeled as in [Pu11]. Optical power at the receiver is -10dBm and a post-detection RRC filter with roll-off 1 and 10GHz bandwidth is applied on the photo-detected signal; thermal noise from photodetectors is also included. BER measurements are performed by means of symbol-by-symbol hard threshold detection; no DSP is performed on the acquired waveforms prior to BER measures, except timing recovery to find the optimum decision point within the symbol time, as well as the optimum decision thresholds.

In the first set of experiments, we focus on analyzing one single PAM4 signal (labeled 4 in Fig. 6-1) that is simulated to drift right with steps of 200 MHz up to 5 GHz; the BER degradation was plotted in Fig. 6-8, where after 3 GHz of LD, the BER starts to increase sharply due to detuning w.r.t. the receiver optical filter. Selecting the spectra of PAM4 of the simulation, the signal identification and classification algorithm perfectly identifies both signals and matches them to two existing lightpaths; the algorithm classifies as *normal* signals when both are still in their expected spectral window, whereas it identifies the signal that has moved out of its expected spectral window as *outOfRange*.

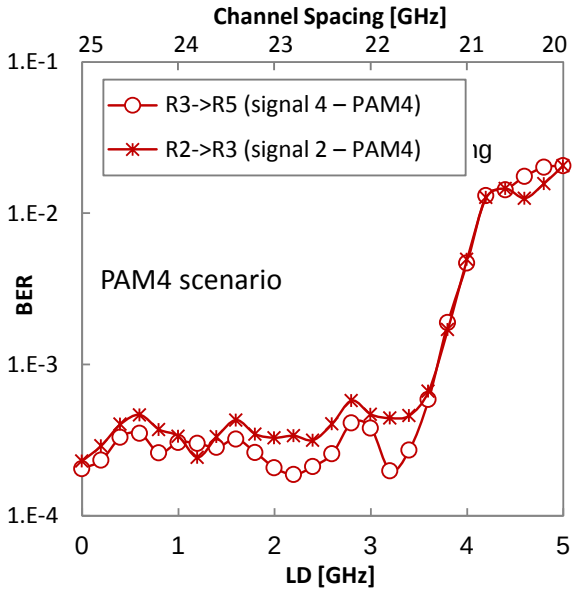


Fig. 6-8. BER degradation vs LD and/or channel spacing for PAM4 systems

Additionally, it is important to investigate whether our proposed LD estimation algorithms work well with acquisitions of OSAs of different resolutions. Therefore, we carried out LD estimation considering OSAs of different granularities; the results are plotted in Fig. 6-9. The results are illustrated for the three approaches described in the previous section. As shown, the estimation accuracy of the approaches relying on the residuals and super features is almost perfect regardless of the OSA resolution (residual based approach performs even slightly better),

while the accuracy of the approach relying on the individual feature degrades as OSA of coarser granularity is considered.

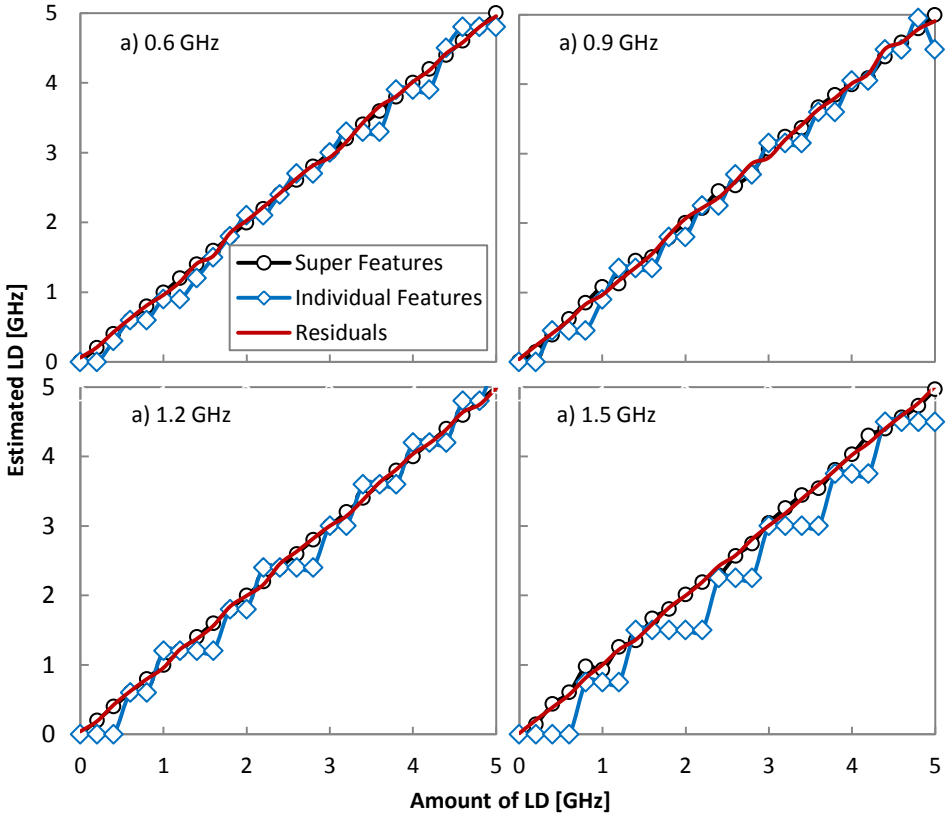


Fig. 6-9. The tracking accuracy of the PAM4 optical spectrum

In the second set of experiments, three PAM4 signals spaced 25 GHz apart, are modeled. Similar to Fig. 6-1, the CF of the central one (signal 2) is detuned towards the channel on the right (signal 3) with steps of 200 MHz up to the point their spacing becomes 20 GHz. Now, we are interested in evaluating whether the proposed approach enables reducing the channel spacing to reduce the margins.

In this case, LD estimation for signal 2 works as accurate as in the previous experiment for signal 4. However, when the individual analysis procedure is applied to analyze signal 3 we observed that the accuracy was poor; the results are presented in Fig. 6-10a, where the estimation error is plotted as a function of the spacing between the signals. These results are as a consequence of part of signal 2 overlapping signal 3, which for coarse-granular OSA produces a significant loss of accuracy. Fig. 6-10b presents the results obtained when contextual analysis is applied. After finding that signal 2 presents some amount of LD, the contextual approach generates an estimated signal that takes into account such fact. This contextual expected signal generates residuals that counteract the effects of signal overlapping and allows preserving the accuracy observed in the first set of experiments. In fact, is that extraordinary accuracy what enables reducing channel

spacing, as any slight LD can be detected and procedures for laser retuning triggered.

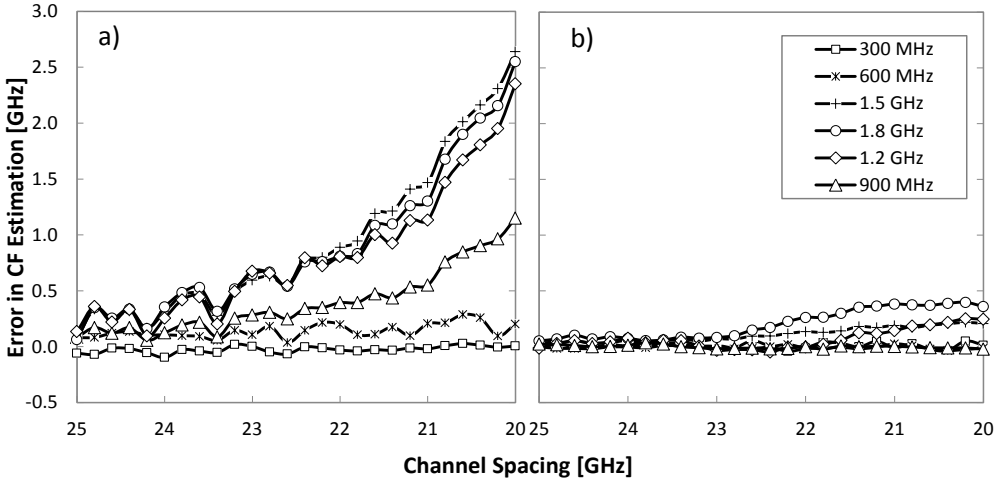


Fig. 6-10. The tracking accuracy of the PAM4 signal with a) individual signal analysis, and b) contextual analysis

6.4.2 QPSK Scenario

In this second scenario, we apply the algorithms to a set of experimental measurements of QPSK signals. We setup an experimental testbed where two neighboring 100 Gb/s signals (labeled *s1* and *s2*) were launched. Signal *s1* was generated using an experimental system, while signal *s2* was generated by a commercial system. Signal *s2* was considered to operate properly, while *s1* is forced to move toward the neighboring one at 1 GHz steps from an initial 50 GHz spacing between signals (Fig. 6-11), simulating a laser drift failure.

Due to the filter-less characteristics of the network, signal *s2* is affected after a certain the amount of laser drift. Fig. 6-12 shows the pre-FEC BER of signal *s2* as a function of channel spacing between the two signals. A sequence of C-band spectra was acquired using a commercial OSA with 100 MHz resolution; a number of captures with coarser resolution, from 300 MHz to 3 GHz, were subsequently generated from the original capture to analyze the impact of the resolution on the accuracy of the proposed algorithms.

Similar to the PAM4 scenario, selecting the spectra in Fig. 6-11, the signal classification algorithm perfectly identifies both signals and matches them to two existing lightpaths; the algorithm classifies as *normal* signals when it analyzes the spectrum in Fig. 6-11a, whereas it identifies *s1* as *outOfRange* when analyzes that in Fig. 6-11b. Nonetheless, it is worth studying the accuracy of signals' detection vs. OSA resolution. To that end, we emulated 5000 different lightpath frequency ranges for every spectrum capture with no overlap.

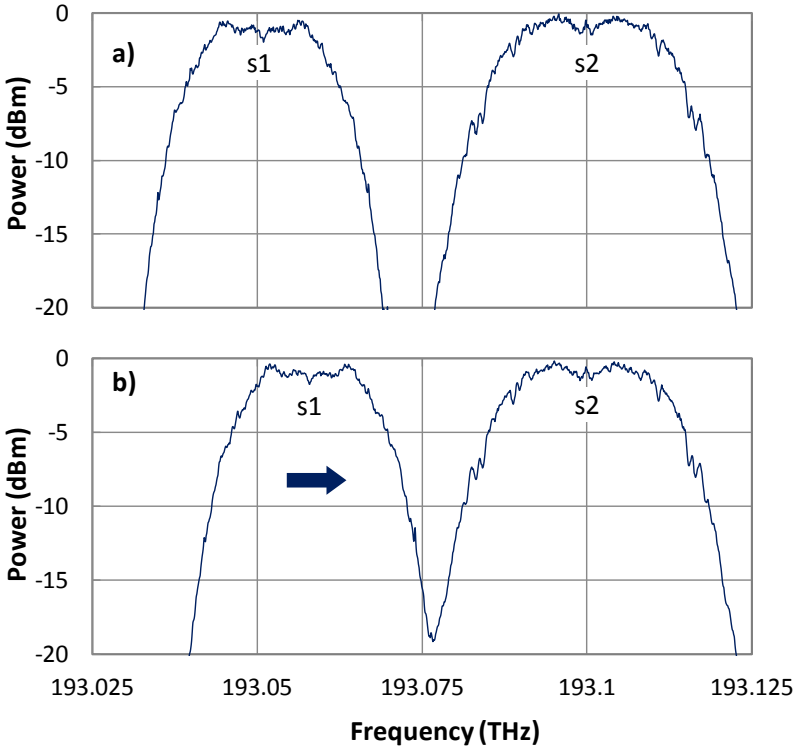


Fig. 6-11. Optical spectrum under normal conditions (a) and when LD causes overlapping (b)

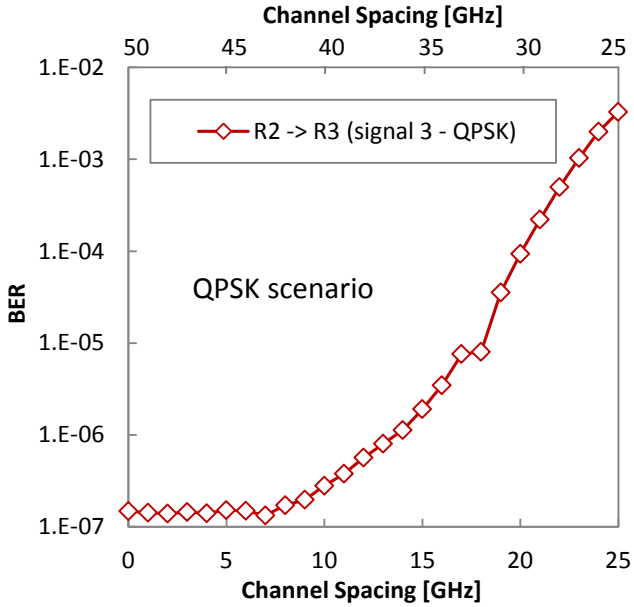


Fig. 6-12. BER degradation vs LD and/or channel spacing for a QPSK system

The accuracy results of $f_{1(\text{-3dB})}$ / $f_{2(\text{-3dB})}$ computation vs. OSA resolution are reported in Fig. 6-13a. The finer the OSA resolution the lower the error in points' computation, which impacts on signal identification.

Let us now study the accuracy of signal overlapping detection as a function of the OSA resolution. In this case, we track $f_{1(\text{-3dB})}$ and $f_{2(\text{-3dB})}$ to infer when both signals begin to overlap. Fig. 6-13b shows the results of the detection; with 300 MHz and 600 MHZ OSA resolution, the overlap is perfectly detected; the inner graph inside Fig. 6-13b shows how the sudden change in $f_{1(\text{-3dB})}$ of signal *s2* allows detecting the overlap. When OSA resolution is up to 2.1 GHz, 1 GHz of effective overlap is needed to detect it, whereas the overlap is not detected for coarser OSA resolutions. Note that, we defined effective overlap as the spectral distance between the optical spectrum traces of two signals at power – 20dB.

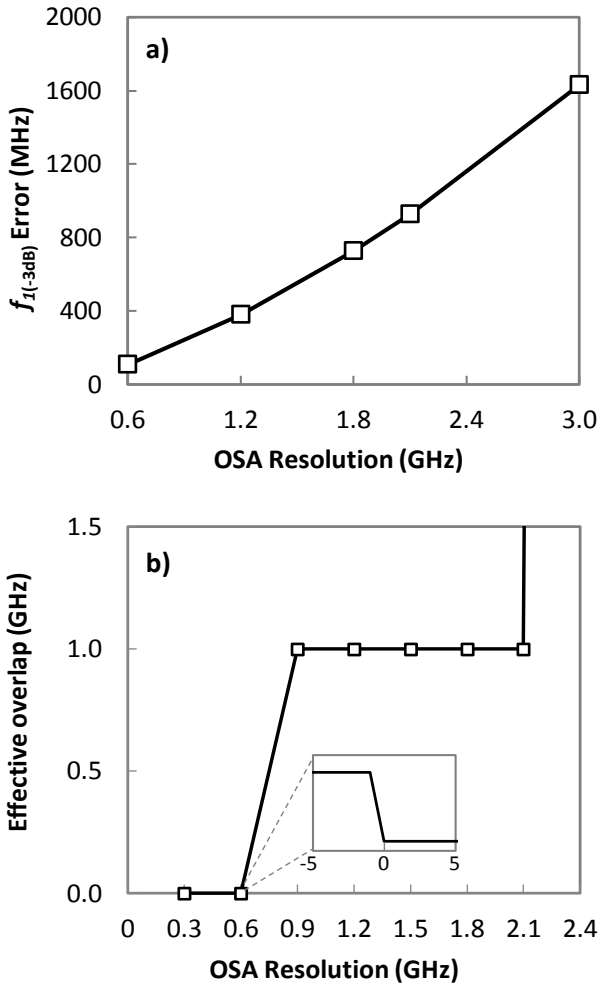


Fig. 6-13. $f_{1(\text{-3dB})}$ computation vs OSA resolution(a) and overlap detection vs OSA resolution (b).

Ultimately, we report the accuracy of the three LD estimation approaches to the case of QPSK-modulated signals. The results are illustrated in Fig. 6-14. Similar to the PAM4 case, the tracking accuracy of the approaches that are based on residuals and super features is almost perfect. However, the accuracy of the approach based on individual features becomes worse as the coarser OSA resolution is considered.

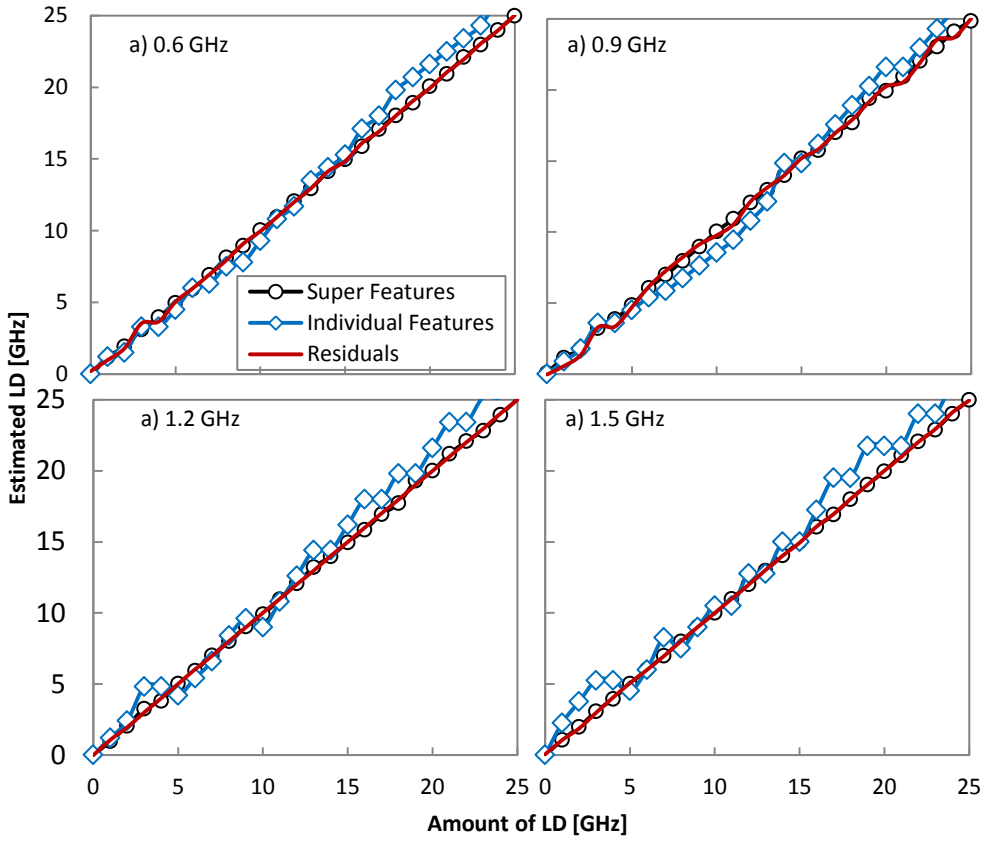


Fig. 6-14. The tracking accuracy of the QPSK optical spectrum

6.5 Concluding Remarks

In this chapter, a data analytics based surveillance system exploiting the optical spectra of lightpaths, to detect anomalies and failures, has been proposed. The proposed surveillance system comprises of signal identification, classification, and tracking modules. The main purpose of signal identification and classification modules is to check whether all the lightpaths are placed in their expected spectral windows and if not, to notify the controller. The optical signal tracking modules, on the other hand, is in charge of monitoring movement of properly operating lightpath, properly, due to improper operation or misconfiguration of a transponder. Optical signal tracking module can be used: i) to avoid the misalignment of a transmitter with the optical filter of its corresponding receiver in a direct detection system, and ii) to predict whether two neighboring lightpath may overlapping due to the failure of a failure. It can be also used to realize system operating with low channel spacing contributing to low-margin operation of the network. The performance of all the approaches have been demonstrated by applying to a set of measurement collected through simulation in a PAM4 direct

detection system and a set of measurements acquired in an experimental testbed of 100G QPSK signals.

Optical spectrum monitoring is one of the promising solutions for detecting soft-failures; however, there are other monitoring parameters allowing the improvement of the performance of optical networks. One of the rather unexplored metrics is SOP. In the next chapter, we present a ML-based algorithm, which use the SOP to provide pre-FEC BER estimation.

Chapter 7

SOP Monitoring

Previous chapters focused on utilizing optical spectrum to detect and identify failures in optical networks, which eventually contributes to optical network performance enhancement. However, such approaches are limited to the cases where the optical spectrum of the lightpaths are informative and most importantly can be extracted. In this chapter, we are going to present a solution that contributes to the automation of optical transmission systems. More specifically, an autonomic transmission agent based on machine-learning is proposed for excessive BER prediction resulting from real-time analysis of SOP. The accuracy and speed of the agent enables its use for local and remote transceiver reconfiguration.

7.1 Motivation

As the coherent technology is introduced in metro networks, energy and cost constraints become predominant [Ku17]. A solution is to reduce the number of bits of Analog-to-Digital Converters (ADC), which are mandatory for coherent technology, to only 4-5 bits instead of 8 bits typically used in long haul networks. This causes a penalty in terms of performance [Ch16] and, thus, the resulting BER previous the Forward Error Correction (FEC) is more sensitive to physical fluctuations, such as the SOP rotation. Moreover, soft-decision FEC mechanisms, such as iterative Low Density Parity Check (LDPC) decoding allow keeping maximum QoT performance [Do12]. Therefore, an adaptive retuning of the FEC mechanism may bring huge interest in this regard.

The motivation behind implementing autonomic transmission in metro networks is supported by the preliminary numerical analysis in Fig. 7-1; the potential benefits in terms of energy savings that could be achieved by reducing the number of FEC iterations is represented in Fig. 7-1, based on the analysis performed in [Do12]. Considering 10 iterations as the static FEC configuration to ensure reliable transmission, power consumption savings above 50% can be obtained by a

reduction of 4 iterations. On the other hand, FEC limit, defined as the maximum pre-FEC BER that can be corrected with no post-FEC errors, can be also tuned w.r.t number of FEC iterations. For this reason, there is clearly a need to implement an intelligent autonomic system that keeps the number of iterations as small as possible, but with enough large to guarantee that the pre-FEC BER limit is not violated and therefore, no post-FEC errors are observed.

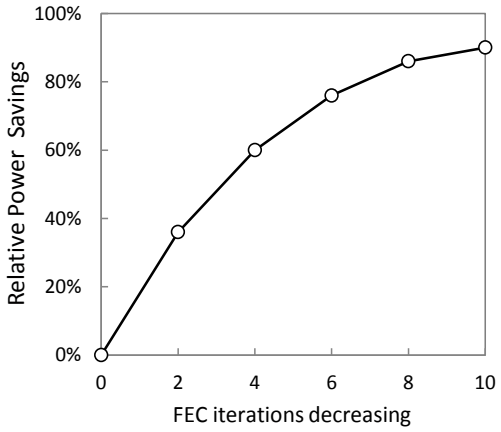


Fig. 7-1. Energy saving with tuning FEC mechanism

7.2 Autonomic Transmission

We propose an autonomic transmission solution, as an extension for current coherent based transceivers, focuses on anticipating pre-FEC BER degradation of individual lightpaths by monitoring in real time the Stokes parameters representing the SOP. In this work, we target two alternative use-cases where the SOP is monitored: *i*) the anticipation of BER degradation can trigger a reconfiguration of the local receiver by increasing the number of FEC decoding iterations; *ii*) the remote transmitter can also be reconfigured by changing, e.g., the FEC coding rate, modulation format or baud-rate. While the local reconfiguration usually takes short times, BER degradation can be anticipated just few milliseconds ahead. In contrast, remote reconfiguration requires longer times since it entails the exchange of messages between the end-nodes through the control plane, so BER degradation needs to be anticipated with more time, e.g., 100ms.

Targeting the above, we propose an Autonomic Transmission Agent (ATA) to lower the impact of physical layer degradations. Considering a point-to-point connection shown in Fig. 7-2, the ATA receives monitoring data from the receiver (labelled 1), analyzes the data, and provides guidelines and operational decisions for the local receiver (2a) or the remote transmitter through the node controller (2b). Sophisticated ML tools are key building blocks of the ATA.

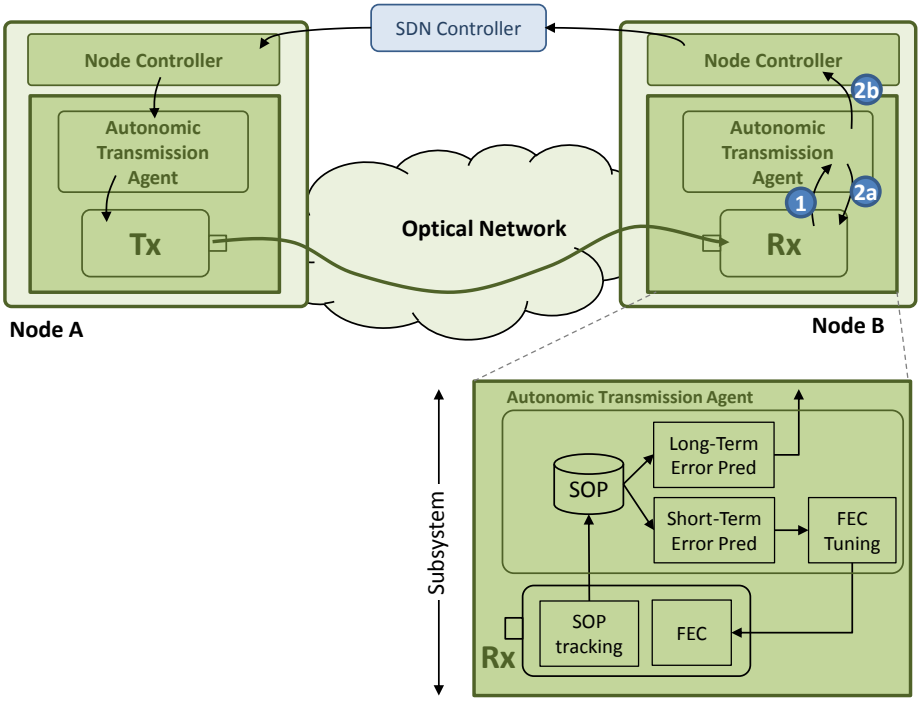


Fig. 7-2. Autonomic transmission and detail of receiver side

7.3 SOP-based Pre-FEC BER Prediction

In this section, we present a set of algorithms for the ATA that continuously receives SOP monitored by the receiver and provides short-term and long-term BER predictions. Short and long-term predictions can be interpreted as two complementary phases of prediction. In the first phase (for the next tens of ms), the algorithm estimates the BER after crossing a certain threshold. In the second phase (for the next hundreds of ms), the algorithm provides a long-term anticipation of the evolution of BER.

Fig. 7-3 plots the evolution of the Stokes parameters and the BER for a case where an unexpected event affecting the proper transmission of a lightpath occurs at time t_0 (see [Bo17] for more examples). Both SOP and BER start fluctuating after t_0, revealing some correlation between them. In view of that correlation, we propose a predictive model that can be used at any time t to forecast the expected BER for a given future time $t+\delta$ based on the SOP data measured in the past time window $[t-w, t]$. The model can be used for both long and short-term BER prediction by training it with the proper δ for the targeted prediction.

The predictive model, based on the naïve Bayes approach (see Section 2.4.3), aims at mapping the SOP evolution within the time window $[t-w, t]$ with the probability that the BER takes a given value at time $t+\delta$. In practice, the model can be implemented as a hash table, where each key encodes a specific SOP evolution

pattern and contains a vector of probabilities of all the BER values observed for that SOP evolution.

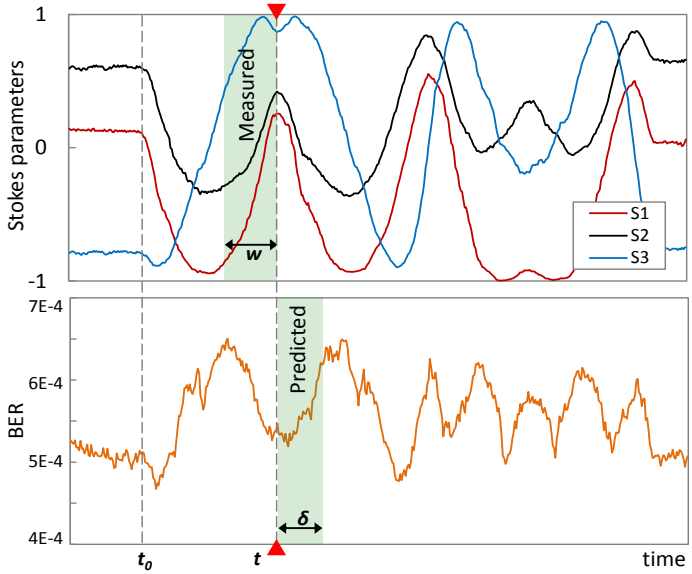


Fig. 7-3. SOP-based pre-FEC BER prediction

The advantage of this probabilistic model is three-fold:

- it allows considering any possible complex correlation between SOP and BER, without any prior assumption on that;
- *ii*) using probabilities increases its flexibility for different cases. For short-term prediction, a simple weighted average BER can be obtained as a good BER estimator. However, since large δ impact negatively on accuracy, the probability that BER exceeds a given threshold becomes a more robust prediction outcome; and
- *iii*) its computational cost is almost negligible, which allows very fast operation.

The algorithm in Table 7-1 is used for model training from a dataset $D_{w\delta}$ containing SOP and BER monitoring measurements for a specific pair w, δ. Let us consider that each row in $D_{w\delta}$ contains a number of Stokes values for a given window [t-w, t] and the BER measured at the target time t+δ. After initializing the empty model F, the training procedure starts by summarizing each row into 6 relevant features (two per Stokes parameter) related to SOP evolution: the Stokes parameters measured at time t and their trend within the window (lines 1-2 in Table 7-1). Those continuous values are then discretized into p segments of equal length (line 3). For example, for a given Stoke parameter (whose continuous range is [-1, 1]) and p=20, segment 1 will contain those rows in the range [-1,-0.9), segment 2 those in the range [-0.9,-0.8), and so on. Similarly, BER is also discretized into q segments of equal length (line 4). Note that discretization of SOP and BER

measures returns also continuous ranges (r), which are needed to translate continuous values to segments and vice-versa.

Table 7-1: Training algorithm

INPUT: $D_{w\delta}$, p, q
OUT: $model_{w\delta}$

```
1:  F ← Ø
2:  Xc ← computeFeatures(D_wδ.SOP)
3:  X, r_SOP ← discretize(Xc, p)
4:  Y, r_BER ← discretize(D_wδ.BER, q)
5:  for i in 1..length(X) do
6:     key ← labeling(X[i])
7:     if F[key][Y[i]] == Ø then
8:        F[key][Y[i]] ← 0
9:     F[key][Y[i]]++
10: for key in F.keys() do
11:    F[key][*]←F[key][*] / sum(F[key][*])
12: return {F, r_SOP, r_BER }
```

After obtaining discrete SOP features and BER, every row is firstly mapped to a label according discrete SOP features (lines 5-6); a row whose discrete features take the values 1, 5, 13, 20, 17, and 5, is labelled as '*010513201705*'. Then, the count of rows of that SOP label with the BER segment is incremented (lines 7-9). This procedure ends with a table that contains the count of each SOP label and BER segment, which is easily converted into an empirical probability by dividing cell values over the sum of the rows defined by keys (lines 10-11).

7.4 Illustrative Results

The accuracy of the proposed algorithms in predicting pre-FEC BER variations is evaluated in this section. The SOP and the corresponding dataset have been collected from an experimental setup that includes a QPSK optical coherent transmission, implemented on a real-time Field Programmable Gate Array (FPGA) (see [Bo17] for details) using a low resolution 5-bit ADC, along 50km of SMF fiber. Mechanical events are caused by a programmable robot arm twisting the fiber.

SOP monitoring samples, extracted from the de-multiplexed polarization steps in the digital signal processing are gathered at a rate of 3,600 samples per second. Every experiment is triggered by a rotation speed threshold and contains 2000 samples before the triggering event and 14000 samples after it. At every time interval, the number of errors is evaluated by counting the difference between the expected pseudorandom sequence and the decoded one (by step of 32 bits) on each

tributary. The number of errors is then computed by considering the baud rate, the polarization multiplexed factor, and the QPSK number of bits.

500 different experiments were carried out. Out of these experiments, we selected ~100 experiments for training and testing the proposed algorithm in which, due to mechanical stress, the BER significantly increases beyond proper operation period measurements. For the BER prediction, we use the relative prediction error (RPE), defined as,

$$RPE = \frac{predictedBER - realBER}{realBER} \tag{7-1}$$

as a metric to evaluate the prediction accuracy of the algorithm. We aim at accurately predicting those higher values of BER, i.e. when it exceeds the mean plus 3 times the standard deviation measured during the proper operation period.

The obtained results are shown in Fig. 7-4, where the mean and maximum values of RPE are presented for a range δ. High prediction accuracy for the 50ms short-term target is shown, with only 3% mean and about 7% of max RPE. However, the accuracy decreases as longer prediction time windows are considered. To complement the above results, Fig. 7-5 shows BER prediction using a model for δ=10ms for two different experiments; one can observe how algorithm accurately follows the evolution of BER, regardless of SOP fluctuations.

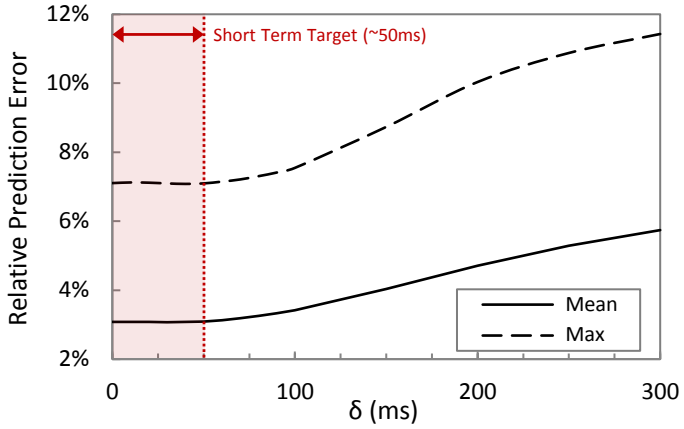


Fig. 7-4. Prediction accuracy

As the accuracy predicting the BER decreases for longer time windows, in the second phase we focus on anticipating whether the BER will exceed a particular threshold and therefore, the accuracy of the algorithm will be evaluated in terms of BER threshold violation. In order to setup this threshold, the mean and standard deviation of errors during the operation phase was computed for every experimental measurement and then, the threshold was set to the mean plus 6 times the standard deviation.

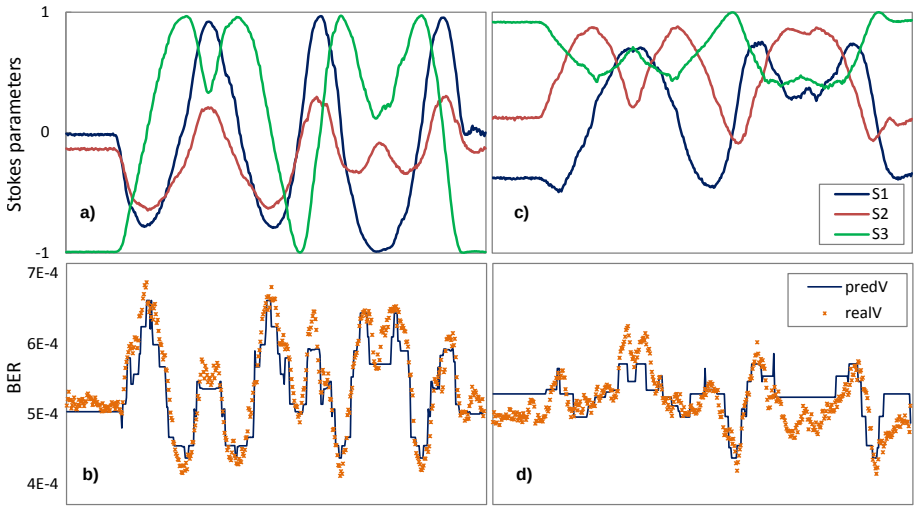


Fig. 7-5. BER estimation of two different SOP profiles

We first trained the model with different w ranging from 0 to 200 ms; the obtained results are presented in terms of the threshold violation anticipation time as a function of w (Fig. 7-6). In view of the results, we can conclude that w=150ms, i.e., last 540 SOP samples (6 kB storage), allows the best overall accuracy and allows anticipating threshold violation up to 45% times sooner than other window sizes (dashed line). These results allow validating the proposed methodology for both use cases and additionally, to highlight that different models are needed for different prediction time scopes.

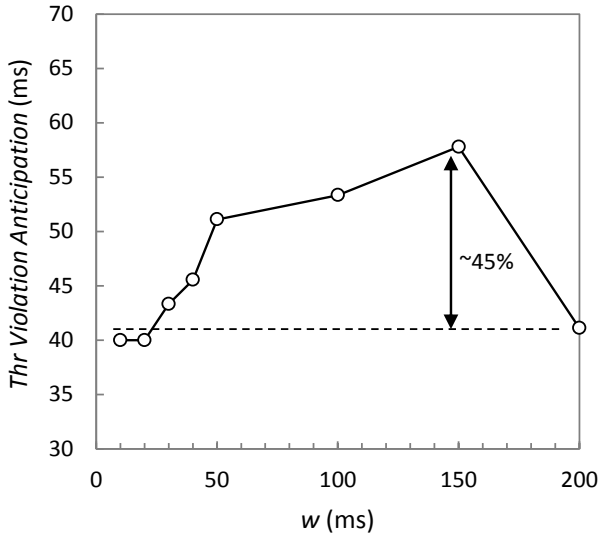


Fig. 7-6. Threshold violation anticipation vs. w

With that optimum w configuration, the analysis of how long δ can be to guarantee an accurate threshold violation detection is presented in Fig. 7-7, which plots the actual threshold anticipation time as a function of δ. Although models could

potentially target anticipation time equal to δ (dashed line), this becomes slightly smaller since increasing δ makes it more difficult to detect threshold violation when it happens and consequently, this delays the anticipated detection. However, 80ms of actual threshold anticipation can be achieved when δ=100ms. After that, detection time drops and therefore, δ=100ms is proposed for long-term models to anticipate threshold violation and trigger some reconfiguration procedure at transmitter location.

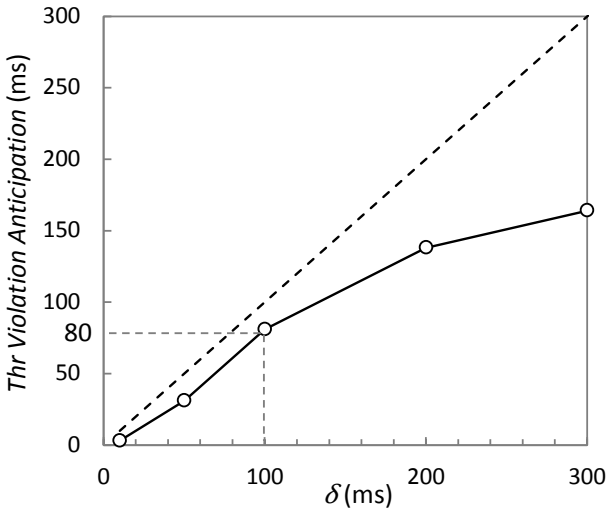


Fig. 7-7. Threshold violation anticipation vs. δ

7.5 Conclusions

An autonomic transmission agent has been proposed to enhance the performance of optical transmission systems. The agent benefits from extremely low complexity, minimal storage, and speed to leave sufficient time for local and remote transceiver reconfiguration.

In the last four chapters we have presented the conducted research focusing on the development of ML based solutions, which allow EONs to either sustain their proper operation or take prompt action in case of unexpected events. In the rest of this PhD thesis, we focus on the performance evaluation of SDM networks utilizing different SDM switching strategies and fiber types. Throughout these investigations, we aim at revealing the benefits and drawbacks of different switching strategies and fiber types for the realization of SDM networks.

Chapter 8

Performance Comparison of Different SDM Switching Schemes

In this chapter, we move forward and focus on the performance evaluation of several future SDM networks and compare that of different SDM network realizations exploiting various switching strategies. As introduced in Section 2.5.2, SDM switching architectures have been proposed to provide switching capability for networks based on spatially integrated optical fibers, as well as to reduce the implementation cost. These architectures rely on the following three switching paradigms, furnishing different degrees of spectral and spatial switching granularity: Independent Switching (Ind-Sw), which offers full spatial-spectral flexibility; Joint Switching (J-Sw), which treats all spatial modes as a single entity; and Fractional Joint Switching (FrJ-Sw), whereby sub-groups of spatial modes are switched together as independent units. The last two paradigms can be categorized as Spatial Group Switching (SG-Sw) solutions since the spatial resources are switched in groups rather than independently, as in the case of Ind-Sw. The contributions reported in this chapter can be categorized into two main folds:

- The impact of traffic profile on the performance of the SDM switching paradigms listed above is investigated in terms of blocking probability. We show that the performance of these switching strategies are traffic-dependent and reveal the circumstances under which each switching strategy can show its best performance.
- The impact of spatial and spectral switching granularities on the performance of the SDM switching paradigms is presented in terms of spectral utilization and data occupancy. We consider two WSS technologies for the handling of the SDM switching paradigms: 1) the current WSS realization, and 2) WSS technology with a factor-two resolution improvement. We show that the performance of all switching paradigms

converges as the size of the traffic demands increases, but finer spatial and spectral granularity can lead to significant performance improvement of SG-Sw schemes for small traffic demands.

8.1 Impact of Traffic Profile on the Performance of SDM Networks

In this section, we thoroughly evaluate in a dynamic scenario the impact of various traffic profiles and dimensioning approaches on the performance of SDM switching paradigms. We show that the performance of SDM switching paradigms is highly dependent on the traffic profile. While J-Sw shows similar performance as Ind-Sw for large demands, it presents a worse performance when the network is fed by a large number of small demands. However, we show that the performance of the J-Sw case can improve substantially when finer spectral granularity is considered.

8.1.1 Simulation Environment and Assumptions

In this study, we use the Telefónica Spanish national network. It comprises 30 nodes (average nodal degree 3.7, max. 5), 14 of which support add/drop (A/D) capabilities, as well as 56 links with an average length of 148 km. In order to have a fair comparison among the three SDM switching paradigms, regardless of any transmission medium related performance constraints, bundles of 12 SMFs (i.e., BuSMFs) are considered for all links in the network. As discussed in Section 2.5.3, we define BuSMFs as a set of individual SMFs, which are packed together in a single strand of cable. Moreover, based on the network characteristics and the related performance evaluation studies [Pa13], DP-8QAM at 32-Gbaud was chosen as the modulation format offering the best compromise between optical reach and spectral efficiency. For these studies, let us assume 50 GHz channel spacing.

According to the above and considering an available spectrum per fiber equal to 4.8 THz (C-band) on the ITU-T 12.5 GHz grid, discrete event simulation studies were carried out for performance evaluation. More specifically, the RSSA problem is solved with a k-shortest path (k = 3) and spatial Super-Channel (Sp-Ch) allocation algorithm, that follows the first-fit strategy, starting with the shortest computed path. The load generation followed a Poisson distribution process. Traffic demands for each source-destination pair were generated randomly following a normal distribution with mean μ and standard deviation σ over the range of study, namely 50 Gb/s to 2.25 Tb/s. Blocking Probability (BP), defined as the amount of blocked traffic divided by the total amount of traffic fed to the network, was used as a quantitative performance measure.

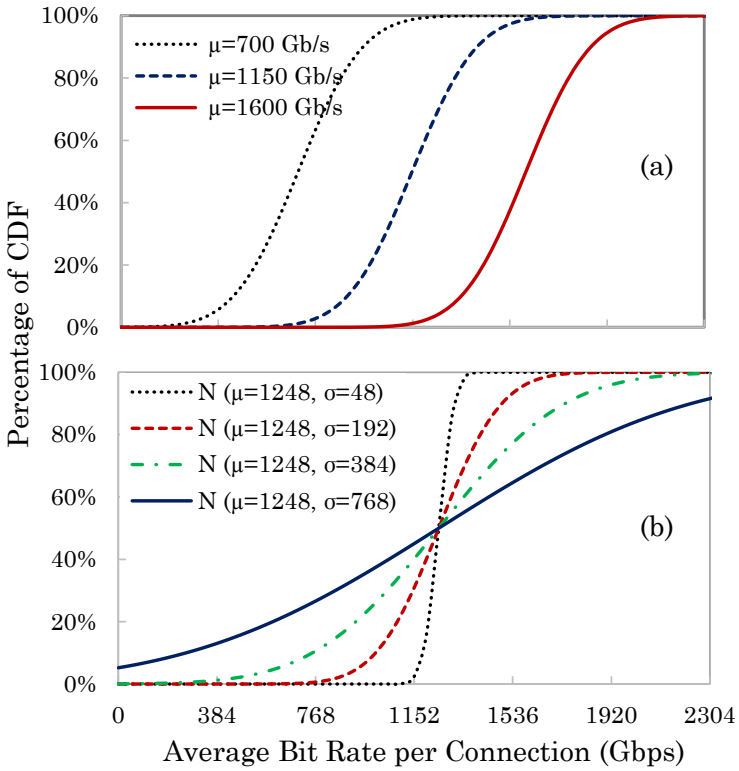


Fig. 8-1. CDFs of the assumed traffic profiles with (a) fixed σ=200 Gb/s and μ=700, 1150, and 1600 Gb/s, (b) fixed μ=1248 Gb/s, and σ=48, 96, 192, 384, 768 Gb/s.

8.1.2 Illustrative Results

In the first part of the study, we consider three traffic profiles corresponding to the three different cases: small, large and medium-size demands. The distribution of demands for the three different mean demand values and fixed σ is shown in Fig. 8-1(a), while the effect of the deviated demands over a fixed mean value is shown in Fig. 8-1(b). The lower and upper mean values were chosen according to the following: a) for μ=700 Gb/s, 98% of demands requires less than half of the 12 spatial dimensions (i.e. BuSMFs in our case study) to be allocated. b) for μ=1600 Gb/s, we have large aggregated demands that result in more than 98% of them requiring more than half of the 12 spatial dimensions to be allocated. The three traffic profiles of Fig. 8-1(a) are used to obtain the results shown in Fig. 8-2. It is noted that in all cases traffic dimensioning is realized by varying the number of live connections per A/D node.

For high mean traffic demands (Fig. 8-2(a)), the three switching paradigms show the same performance. Since most demands require more than half of the spatial resources, the unutilized resources of FrJ-Sw and Ind-Sw cases cannot be allocated thus leading to the same results as the J-Sw case. For a traffic profile with diverse and relatively medium-size demands (Fig. 8-2(b)), FrJ-Sw and Ind-Sw start

performing better than J-Sw in terms of BP, since now part of the incoming demands that require less than 6 spatial dimensions to be allocated can fit within the free spatial resources that FrJ-Sw and Ind-Sw enable them to use. For small mean traffic demands (Fig. 8-2(c)) the performance difference between J-Sw and FrJ-/Ind-Sw is more pronounced since the allocation options in space dimension increase and small demands can fit in available spatial slots.

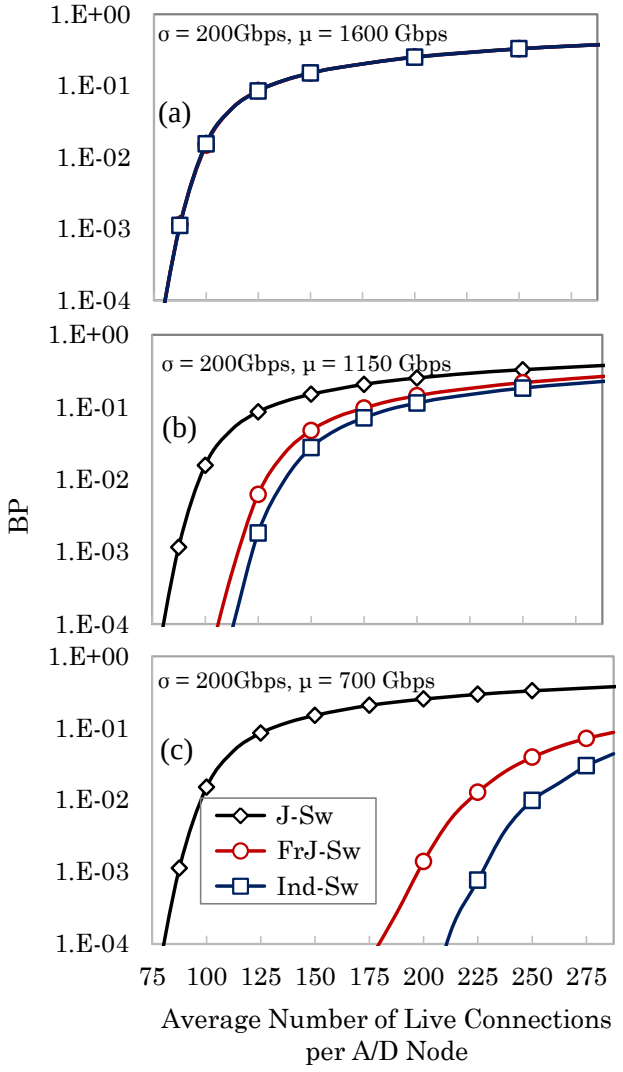


Fig. 8-2. BP in terms of average number of live connection per A/D nodes

Previous discussions on Fig. 8-2 strongly suggest that the optimal switching paradigm for an SDM network, in fact, depends on the nature of its traffic, specifically whenever there is a prevalence of relatively small or large demands. However, since most of the demands in the core networks are aggregated traffics, J-Sw would be a suitable choice, considering its cost benefit as it will be presented in Section 10.1.

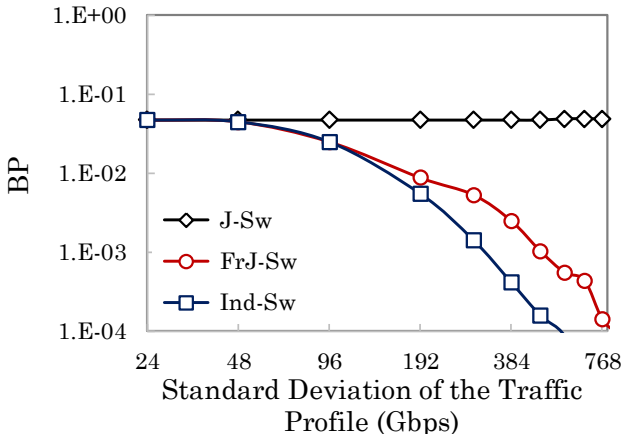


Fig. 8-3. BP in terms of standard deviation when μ=1248 Gb/s and average number of live connections per A/D node is 112.

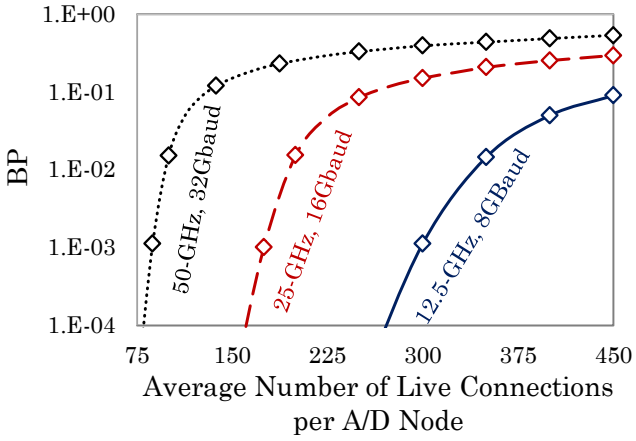


Fig. 8-4. BP w.r.t an average number of live connections per A/D node with μ=350 Gb/s and σ=200 Gb/s for J-Sw assuming three levels of spectral granularities.

In order to see the impact of traffic diversity on the performance of SDM switching paradigms, a complementary set of simulations is carried out, where the traffic dimensioning is done by varying σ and keeping μ and the number of live connections fixed (Fig. 8-1(b)). Note that, in this study, the total offered load to the network during the whole range of the simulation is fixed to 14×112×1248 Gb/s = ~1.95 Pb/s. Results plotted in Fig. 8-3 show that at the beginning the performance of three SDM switching is the same (similar to Fig. 8-2(a)), and by increasing σ, which is equivalent to the diversity level of the traffic profile, Ind-Sw and FrJ-Sw show a remarkable improvement, which justifies their suitability for networks with a high level of diversity in their traffic profile.

As shown in Fig. 8-2 and Fig. 8-3, J-Sw shows worse performance in case of a large number of relatively small demands. However, since WSSs allow fine switching granularity (i.e. 25 or 12.5GHz), we investigated the performance when the spatial

Sp-Ch occupies 25 or 12.5 GHz spectrum. As shown in Fig. 8-4, the performance improves significantly when switching granularity increases.

In summary, we found that Ind-Sw and FrJ-Sw perform well for networks with high level of traffic diversity, while J-Sw is a favourable option for networks with large demands. However, J-Sw can perform significantly better in high diverse traffic scenarios, if spatial Sp-Chs occupy smaller spectral width.

In the next section, we focus on the impact of different spatial and spectral granularities levels on the performance of SDM switching strategies. We will reveal in more detail how different granularity levels can affect the network-wide performance of SDM networks. Additionally, we explore the network-wide benefits of utilizing finer resolution WSSs for the realization of SDM switching nodes.

8.2 Impact of Spatial and Spectral Granularity on the Performance of SDM Networks

8.2.1 Motivation

In the previous section, we showed that the performance of SG-Sw cases becomes similar to that of Ind-Sw as *the total offered load to the network* (hereinafter referred to as *load*) increases assuming bundles of 12 SMFs. However, SG-Sw cases showed a reduced performance for low values of load. We further showed that SG-Sw paradigms with lower G values perform well for networks with a high level of traffic diversity, while J-Sw is favourable for networks with large demands. Note that in the previous section, Ind-Sw, J-Sw and FrJ-Sw with $G = 3$ were only examined with a fixed Channel Spacing (hereinafter called ChBW) equal to 50 GHz. In this section, in addition to the three levels of spatial switching granularity studied before, we examine FrJ-Sw with G equal to 2, 4, and 6. Specifically, the focus of this section is to compare the performance of the three SDM switching paradigms described before for spatial switching considering various spatial and spectral granularities and employing two WSS technologies for their realization.

Current WSS technology allows allocating channels with 18 GHz guard band according to a 6.25 GHz grid, i.e., 32 GHz can be provisioned on 50 GHz, 25.75 GHz on 43.75 GHz, 19.5 GHz on 37.5 GHz, 13.25 GHz on 31.25, or 7 GHz on 25 GHz. Here, we investigate the impact of the spectral switching granularity on the performance of spatial Sp-Ch switches based on two practical WSS technologies: 1) the current generation WSS realization, and 2) a WSS technology with a factor-two resolution improvement (i.e. requiring 9 GHz for guard band instead of the 18 GHz considered above). A summary of the ChBW values selected for the studies performed in this section and the corresponding clear channel bandwidth that can be allocated with data and switched by current and future (factor-two resolution improvement) WSS realizations is provided in Table 8-1. The impact of ChBW and

WSS resolution in conjunction with the various switching strategies is investigated in this section.

Table 8-1 Values of selected ChBW with the amount of corresponding spectral contents supported by two WSS technologies

Spectral Channel plan [GHz]	50	43.75	37.5	31.25	25
Current generation WSS resolution (clear channel BW and % utilization)	32	25.75	19.5	13.25	7
	64%	59%	52%	42%	28%
Improved resolution WSS technology (clear channel BW and % utilization)	41	34.75	28.5	22.25	16
	82%	79%	76%	71%	64%

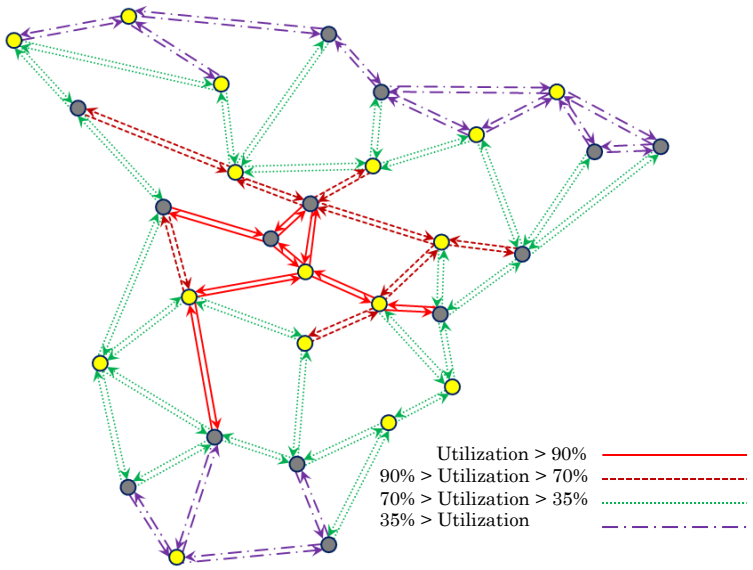


Fig. 8-5. Topology of Telefónica Spanish national network. Nodes with A/D capabilities are illustrated with filled grey circles.

8.2.2 Simulation Environment and Assumptions

The studies reported in this section were performed in an offline planning scenario. The network simulations were carried out over the Telefónica Spanish national network shown in Fig. 8-5. The network is characterized by a symmetric traffic matrix, with 84 connection requests between two subsets of 7 transit nodes, which is significantly heterogeneous. The total amount of traffic (i.e. total amount of all 84 demands) is equal to 10.5 Tb/s estimated for the year 2014, when the smallest demand is ~2 Gb/s, the largest one is ~488 Gb/s, and the mean value of all demands is ~125 Gb/s.

The heterogeneous nature of the traffic matrix is the consequence of connection requests being mostly exchanged between highly populated transit areas (more

than 70% of the total traffic flows are from/to Madrid), where the Internet exchange points are located. This leads to the existence of *hot links* (i.e. links with more than twice the average spectrum utilization per link in the network), *underutilized links* (i.e. links with <1/3 the average spectrum utilization per link in the network) and *moderately utilized links*. The topology of Telefónica's network, in which links are differentiated based on their utilization, is shown in Fig. 8-5.

We limited the network performance study to the case of BuSMFs. This has the advantage that the transmission is not affected by crosstalk between spatial dimensions (fibers), and SDM multiplexers/demultiplexers are not required for component interconnection. Assuming BuSMFs also allows us to make a fair comparison between different switching paradigms, because the BuSMFs are the only type of fiber compatible with all switching paradigms.

We implemented a RSMSA algorithm consisting of a diverse routing computation element (a *k*-fixed alternate shortest path with maximal disjoint links for each source-destination pair [Ei98]) and a resource allocation module. In the resource allocation module, the spatial and spectral resources are assigned to connection requests in the form of spatial Sp-Ch, following a first-fit strategy, starting from the shortest path and the lowest indexed spatial/spectral resource. In order to alleviate the problem posed by the *hot links*, we implemented a load-balancing engine including a *request-breakdown* element, which breaks up connections larger than the capacity of one spatial Sp-Ch, i.e. the number of SMFs in the bundle. The load-balancing engine distributes big connection requests proportionally over *underutilized* and *moderately utilized* links when the shortest paths between end-nodes become very congested. The load-balancing engine is more often used when traffic grows, and more alternate shortest paths between source-destination pairs are required to serve one big connection request.

Ultimately, the selection of the best path and the most adequate spatial/spectral resources to establish a connection is carried out by a simulated annealing meta-heuristic optimization tool equipped with a multi-starting-point generator to avoid local minima, thus yielding nearly optimal global spectrum utilization [Ch11]. Regarding the transmission technology, spatial Sp-Ch was considered employing different modulation formats. The choice of modulation format is limited to DP – Binary Phase Shift Keying (BPSK), QPSK, 8-Quadrature Amplitude Modulation (8QAM) and 16QAM, with maximum transmission reach extracted from [Pa13]. The obtained reach values are presented in Table 8-1.

8.2.3 Illustrative Results

In this section, we first compare the performance of spatial Sp-Ch switching paradigms under different spatial and spectral granularities. We assume bundles of 12 SMFs with 4.8 THz available spectrum per fiber across all links as a near-term SDM solution. The performance evaluation is carried out in terms of load

Table 8-2 Maximum point-to-point transmission reach vs baud rate, for different modulation formats

	Baud rate in GBaud	ChBW in GHz	Reach [km]			
			DP-BPSK	DP-QPSK	DP-8QAM	DP-16QAM
Current WSS technology	32	50	9800	4900	1900	900
	25.75	43.75	10800	5400	2100	1000
	19.5	37.5	12300	6100	2400	1200
	13.25	31.25	15400	7700	3000	1500
	7	25	16650	8300	3250	1650
WSS with factor-2 resolution improvement	41	50	7300	3500	1200	600
	34.75	43.75	7800	3700	1400	700
	28.5	37.5	8300	4000	1500	800
	22.25	31.25	9200	4500	1800	900
	16	25	10500	5200	2000	1000

Since there is a fixed number of demands in the traffic matrix (84 demands), traffic growth is achieved by increasing the size of the demands. In order to perform the studies, we scale the *load* (i.e. total offered load to the network) up to 1 Pb/s, which is equivalent to 10-year total traffic growth, assuming 45% annual traffic increase. For the spatial switching granularity, we consider groups (G) of 1, 2, 3, 4, 6 and 12 fibers out of 12 fibers in the BuSMFs, where G = 1 and 12 correspond to the cases of Ind-Sw and J-Sw, respectively, which offer the finest (i.e., Ind-Sw) and the coarsest (i.e., J-Sw) spatial granularities. Intermediate values represent FrJ-Sw with spatial groups formed of 2-6 SMFs.

Fig. 8-6 shows the results considering present-day WSS technology requiring 18 GHz guard band. Fig. 8-6(a) presents the results for the case of fixed-grid 50 GHz WDM ChBW, in which 32 Gbaud is selected for the contents of each ChBW as it is the maximum baud rate supported by present-day WSS resolution. The average spectrum utilization per link per fiber is used as a quantitative network performance metric. Thus, for example, J-Sw with the coarsest granularity in our studies corresponds to G = 12 and ChBW = 50 GHz, allows for 32 GHz×12=384 GHz for data loading, corresponding to 3072 Gb/s assuming DP-16QAM format.

Note that, DP-16QAM is the most used modulation format, as most of the lightpaths in the Telefónica network can be established within its optical reach limit. Demands smaller than that will occupy the whole spatial-spectral slot, resulting in low data occupancy within the available bandwidth. At the finest granularity of G = 1 and ChBW = 50 GHz, the equivalent minimum spatial-spectral bandwidth slot amounts to 256 Gb/s.

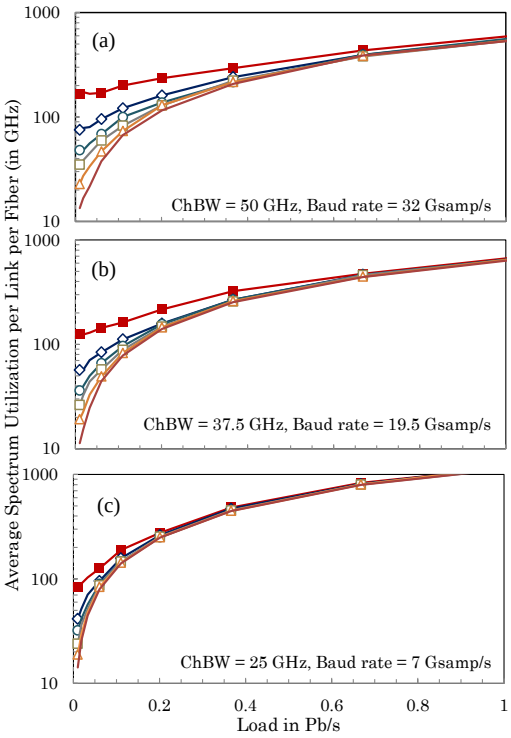


Fig. 8-6. Average spectral utilization considering current WSS technology w.r.t load.

In Fig. 8-6(a), Ind-Sw shows the best performance for all loads, i.e., the lowest utilization for the given traffic load, since it offers the finest granularity. We use it as the benchmark to estimate the unutilized bandwidth due to the grouping of spatial dimensions. The performance of the rest of spatial Sp-Ch switching paradigms is seen to converge to that of Ind-Sw as load increases. This is because when the load increases the spectral-spatial slot becomes comparable to the demand that has to be served and, therefore the amount of unutilized bandwidth due to SG-Sw reduces.

Additionally, we observe that, independently of the load, but more noticeably for smaller loads, the curves for the SG-Sw cases with lower values of G (i.e. finer spatial granularity) are closer to the Ind-Sw curve. This is as a consequence of the higher flexibility that SG-Sw with low G offers to allocate smaller demands in the space dimension.

Current WSS technology with 6.25 GHz assignable spectral slots enables the switching of smaller ChBWs, which allows us to evaluate the impact of spectral switching granularity on the performance of spatial switching paradigms. To carry out this investigation, we repeated the above simulations for 37.5 GHz ChBW at a baud rate of 19.5 Gbaud (Fig. 8-6(b)), and 25 GHz ChBW at 7 Gbaud (Fig. 8-6(c)).

As observed in Fig. 8-6 (b) and (c), for small values of load, all curves show improved performance compared to the 50 GHz ChBW case since finer data loads

can be accommodated per channel. It is noteworthy that the performance of J-Sw (switching paradigm with the coarsest spatial granularity) converges to that of Ind-Sw for smaller load values, as the ChBW decreases, compared to the case of 50 GHz ChBW. For high loads, the spectrum utilization is higher for the case of finer WDM channel grid, as the finite channel guard bands exhibit lower spectral utilization and more channels have to be provisioned to carry the data.

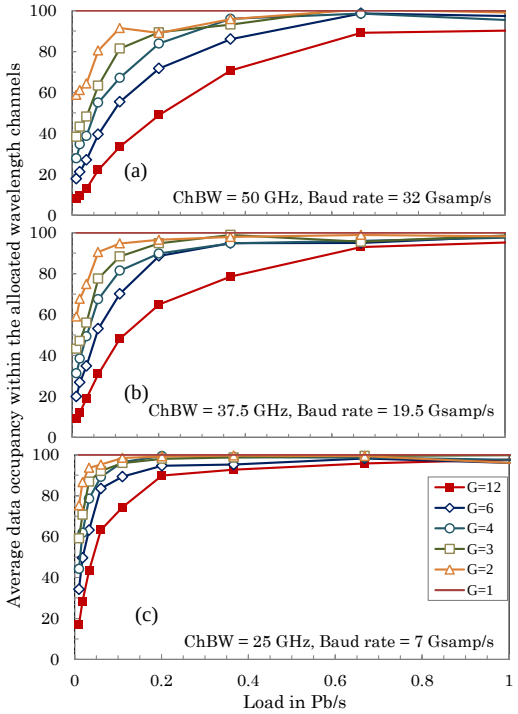


Fig. 8-7. Average data occupancy considering current WSS technology w.r.t load.

Another way to quantify the impact of the switching group size is to consider the average 'data occupancy' within an allocated wavelength channel. The data occupancy metric is defined by the bandwidth required to support the data (i.e. equivalent to the baud rate yet measured in GHz) divided by the available bandwidth for data (which is the clear channel bandwidth multiplied by the group size). Ind-Sw always satisfies 100% data occupancy, whereas J-Sw will have the lowest data occupancy (as low as 1/G). As shown in Fig. 8-7, the amount of data occupancy increases for switching paradigms with finer spatial and spectral granularity.

Additionally, inflection points in the performance of SG-Sw cases are observed at loads of 0.2 and 0.7 Pb/s, respectively, from where the data occupancy of SG-Sw cases increases significantly. In particular, as observed in Fig. 8-7(b), the data occupancy in the case of J-Sw goes above 90% for loads higher than ~0.65 Pb/s with 37.5 GHz ChBW, while it happens for loads above ~1 Pb/s with 50 GHz ChBW. Finally, Fig. 8-7(c) shows a further improvement of the performance of different

SG-Sw paradigms in comparison with the previous two cases (e.g. the data occupancy of J-Sw increases up to 90% at ~0.2 Pb/s).

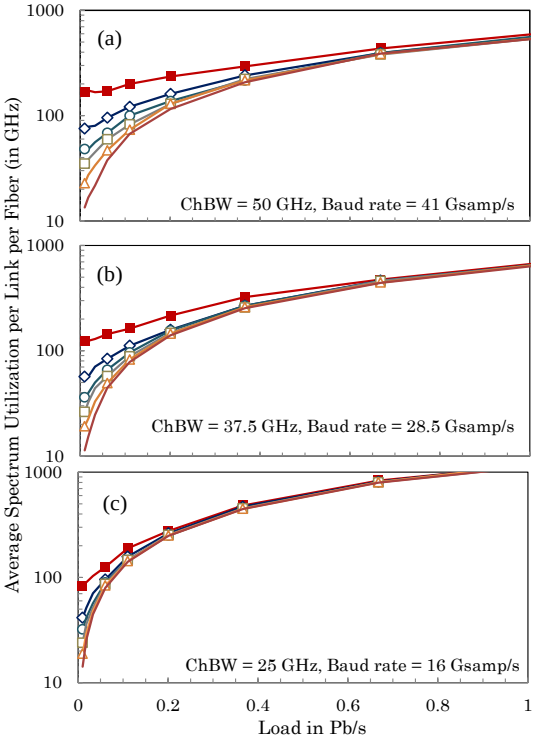


Fig. 8-8. Average spectral utilization considering finer resolution WSS w.r.t load.

Therefore, we can conclude that, for small load values, the utilization of WSSs with finer spectral switching granularity can compensate for the spatial granularity rigidity of SG-Sws. For larger load values, on the other hand, the performance of all switching paradigms is degraded (i.e. the average spectrum utilization increases) as ChBW is decreased. This is due to less efficient utilization of the spectrum arising from a lower amount of occupied spectrum containing actual traffic compared to the required guard band for the WSSs (i.e. 32/50=64% vs. 7/25=28% for ChBW equal to 50 and 25 GHz, respectively).

In order to evaluate the improvement of the SG-Sw performance resulting from the utilization of WSSs with improved resolution, we repeated the above studies for a WSS technology with a factor-two resolution improvement (i.e. requiring a 9 GHz guard band instead of the 18 GHz considered previously). Fig. 8-8 attest that the performance of spatial Sp-Ch switching paradigms can be improved by using spatial switches with finer spatial granularity. Likewise, similar to Fig. 8-6, the use of smaller ChBWs results in performance improvement. Note that, comparing Fig. 8-8 and Fig. 8-6, due to the lower guard band required by the WSS with finer resolution, the performance of switching paradigms is not degraded for large load values, in contrast to the case of current WSS technology with coarser resolution.

However, by comparing Fig. 8-9 and Fig. 8-7, we observe that the amount of data occupancy is less significant when the improved resolution WSSs are in place. This is because one spectral-spatial slot can accommodate more traffic when WSS technology with factor-two resolution improvement is in place compared to the current WSS technology. For example, assuming ChBW equal to 50 GHz and DP-16QAM, the capacity of one spatial-spectral is 72 Gb/s higher (i.e. 328 Gb/s – 256 Gb/s) in the case of improved resolution WSS compared to the current WSS technology.

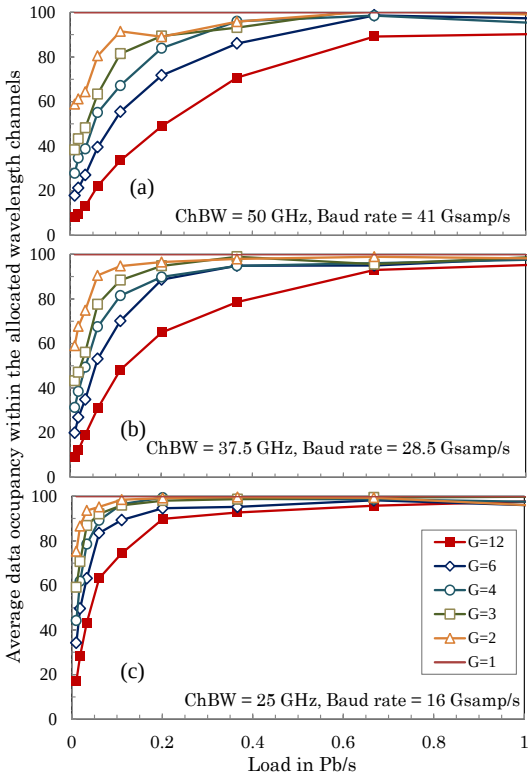


Fig. 8-9. Average data occupancy considering finer resolution WSS w.r.t load.

Note that, even though the data occupancy is lower for WSSs with finer spectral resolution, refining the WSS spectral resolution results in a globally better performance, as is shown in Fig. 8-10. Fig. 8-10 presents the results of a more detailed performance evaluation of J-Sw paradigms for the values of ChBWs indicated in Table 8-1 and for two WSS technologies with coarser and finer resolutions. For the sake of clarity, the results are only shown for J-Sw.

Fig. 8-10(a) shows the average spectrum utilization with the current WSS technology. For small loads (<80 Tb/s), as shown previously, smaller ChBW values lead to better J-Sw performance. However, as traffic increases, smaller ChBW values result in significant performance degradation. Fig. 8-10(b) shows the results when the finer resolution WSS is used. Due to the more efficient utilization of the optical spectrum, smaller ChBW values lead to better performance for loads lower

than 800 Tb/s. Even if the performance of J-Sw with smaller values of ChBW reduces for loads larger than 800 Tb/s, this is remarkably better than in the case of WSSs with coarser resolution. Another important finding is that, for small and large loads, the best J-Sw performance is obtained for the lowest and highest values of ChBWs, respectively. Consequently, ChBW must be adaptable to the load level to achieve a globally optimum spectrum utilization in an SDM network. This highlights the importance of utilizing flex-grid transmission enabled by spectrally flexible ROADMs and bandwidth variable transceivers when SG-Sw paradigms are considered.

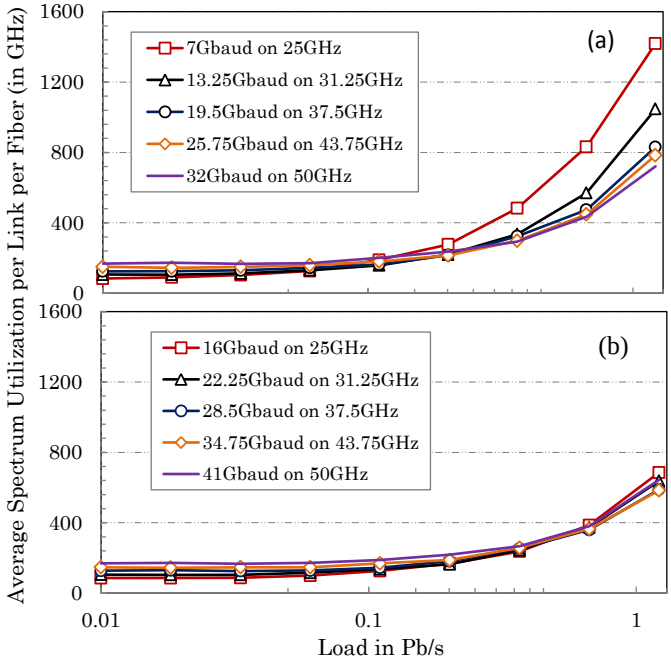


Fig. 8-10. Average spectrum utilization per link per fiber for J-Sw considering: (a) the current WSS technology and (b) finer resolution WSS.

8.3 Conclusions

In this chapter, we have presented a set of comprehensive performance evaluation studies of future SDM networks realized over different switching strategies.

First, we have analyzed the impact of traffic profile on the performance of different SDM switching schemes. It can be concluded that the performance of different SDM switching schemes is traffic-dependent. More specifically, Ind-Sw and FrJ-Sw perform well for networks with a high level of traffic diversity, while J-Sw is a favourable option for networks with large demands, also considering its cost benefit. However, J-Sw can perform significantly better in high diverse traffic scenarios, if spatial Sp-Chs occupy smaller spectral width, which can be switched by WSSs with finer granularity.

Second, we have compared the performance of different SDM switching paradigms, in terms of spectral occupancy under various practically feasible spatial and spectral switching granularities as well as considering the current WSS realization and an improved resolution WSS technology over a network based on BuSMFs. The network-level simulation results showed that the performance of all switching paradigms converges as traffic increases while the switching-related infrastructure cost can be reduced up to 52% in the SG-Sw cases.

Additionally, it was shown that considering the current WSS technology, the utilization of finer spectral switching granularity significantly improves the performance of SG-Sw paradigms for small values of traffic, which correspond to small demands in the traffic matrix. However, as the load increases, the performance of all switching paradigms reduces due to the less efficient utilization of spectrum arising from a lower amount of occupied spectrum containing actual traffic compared to the required guard band for the WSSs. We also showed that, by utilizing WSSs with improved resolution, which requires 50% less guard band, the performance of switching paradigms in the case of large values of traffic can be improved by a factor of two.

It results from our study that, irrespective of the WSS resolution, large values of ChBW are more beneficial for large values of traffic, and consequently spectral switching granularity must be adaptable to the traffic size in order to achieve a globally optimum spectrum utilization in an SDM network, for which spectrally flex-grid ROADMs and bandwidth-variable transceivers are a requirement.

For all the studies presented in this chapter, we have assumed BuSMFs to focus solely on the performance benefits/drawbacks of different switching solutions regardless of the detrimental impact of impairments introduced by different fiber types. However, careful evaluation of the impact of physical layer impairments on the performance of SDM networks utilizing different types of fibers is required to choose the most appropriate fiber type to be deployed in a network. In this regard, in the next chapter, we aim at comparing the network-wide performance of SDM networks realized using different types of fibers.

Chapter 9

SDM Networks Based on SMF Bundles, MCFs, and FMFs

In this chapter, we continue with the study of SDM networks and compare the performance of several SDM network realizations utilizing different types of transmission media. In this regard, we first discuss a tool for the estimation of signal quality degradation in SDM based optical networks. The tool supports SDM networks utilizing FMFs, MCFs, as well as BuSMFs and provides an accurate QoT estimation allowing the network operators and designers to take into account the physical layer impairments while performing network planning analysis.

Considering the QoT tool, as the key contribution of this chapter, we present extensive simulation results comparing the performance of FMFs, MCFs, and BuSMFs in terms of maximum achievable optical transparent reach in a multi-span point-to-point scenario, where the impact of amplifier span length, channel spacing, and number of co-propagating spectral channels is studied. These results are used to perform physical-layer-aware SDM network planning analysis, which allows us to evaluate the impact of reach-limited transmission along different types of fibers in a networking scenario for different sizes of an optical network topology. The results of this study compare different SDM network realizations in terms of average spectrum utilization and the number of in-operation transceivers operating over different modulation formats showcasing the benefits and drawbacks of SDM networks relying on different transmission media.

9.1 Quality of Transmission Estimator

In this section, we describe a QoT estimator for SDM networks that calculates the performance of coherent multi-level phase modulated signals when propagated over an optical uncompensated SDM link (i.e., FMFs, MCFs, and BuSMFs) composed of several amplified spans [Ma17.2]. In this work, we consider FMFs

with M LP spatial modes. Hence, the total mode count including the polarizations is $2M$. The tool supports SDM networks based on both fixed-grid WDM and grid-less multi-carrier multiplexing schemes exploiting Nyquist-WDM and orthogonal frequency division multiplexing. The estimator predicts the QoT of a lightpath to be established in a multi-span point-to-point transmission scenario. Additionally, it is capable of predicting the QoT of new lightpaths to be established in a network, as well as their impact on the existing lightpaths.

The QoT estimation mechanism uses information about the network topology, link characteristics, signal types (Baud rate and modulation format), and lightpaths currently established in the network (i.e., occupied spectral slots on each spatial mode). The estimated QoT is expressed in OSNR per spatial channel and can be translated to Q-factor, BER, error vector magnitude, or optical reach value (in km).

The QoT estimator is developed by using the existing analytical models for each SDM fiber. The OSNR for each SDM fiber type is determined based on analytical models, whose accuracy is verified through experimental evaluations [Po14], [El13], taking into account impairments that are unlikely to be compensated for by DSP at the receiver. So, it is important to have in mind that the analytical models described in the next sections are already available models with demonstrated accuracy through experimental evaluations. The computational complexity of the considered models is low enough to allow the use of the QoT estimator for network planning and operation practices.

9.2 **Analytical Models**

9.2.1 An Analytical Model for FMFs

The closed-form solution used in the QoT estimator for the FMF transmission is the one presented in [El13],[Fe16] and was derived from the Manakov equations for FMFs [Mu13]. A similar approach has also been followed in [Ra16.1]. The model is derived assuming weak-coupling regime, which means the coupling between spatial modes is small or negligible, but the coupling between polarizations is considered strong. As shown in [Ho14.2], [Mu13], [Fe16] the strong coupling regime is expected to reduce the modal dispersion and the nonlinear effects, and, thus, could be used to design optimized fibers where the impact of nonlinear impairments is reduced [Si14.2]. Here, however, working in the weak-coupling regime serves well the current scope of our analysis, as it can be considered as a worst-case scenario for the nonlinear impairments [Mu13].

The Manakov equations can be written in a simplified form [Mu13]:

$$\frac{\partial \vec{A}^{(p)}}{\partial z} + \frac{\alpha}{2}\vec{A}^{(p)} - i(\beta_0^{(p)} - \beta_0^{(1)})\vec{A}^{(p)} + (\beta_1^{(p)} - \beta_1^{(1)})\frac{\partial \vec{A}^{(p)}}{\partial t} + i\frac{\beta_2^{(p)}}{2}\frac{\partial^2 \vec{A}^{(p)}}{\partial t^2} = i(\tilde{\gamma}^{(pppp)}\left|\vec{A}^{(p)}\right|^2 + \sum_{m\neq p}\gamma^{(mmpp)}\left|\vec{A}^{(m)}\right|^2)\vec{A}^{(p)} \quad (9\text{-}1)$$

where $\vec{A}^{(p)}(z,t) = [A_x^{(p)}(z,t), A_y^{(p)}(z,t)]^T$ is the slowly varying envelope of the p-th spatial mode accounting for both x- and y- polarizations, propagating towards the z-direction, in a moving frame referenced to the the group velocity $v_g^{(1)}$ of the 1st spatial mode.

Here, α is the attenuation coefficient and $\beta_0^{(u)}$, $\beta_1^{(u)}$, and $\beta_2^{(u)}$ are the first three terms of the Taylor expansion of the propagation constant for spatial mode p, indicating the mode's phase constant, the inverse group velocity, and the GVD, respectively.

The nonlinear coefficients are defined as [An16]:

$$\tilde{\gamma}^{(pppp)} = \kappa^{(pp)}\gamma^{(pppp)} = \frac{4}{3}\frac{2M}{2M-1}\frac{\omega_0 n_2}{cA_{eff}^{(pppp)}} \quad (9\text{-}2)$$

$$\tilde{\gamma}^{(mmpp)} = \kappa^{(mp)}\gamma^{(mmpp)} = \frac{4}{3}\frac{\omega_0 n_2}{cA_{eff}^{(mmpp)}} \quad (9\text{-}3)$$

where c denotes the speed of light in vacuum, ω_0 is the angular frequency, n_2 is the nonlinear refractive index, M is the number of spatial modes in the FMF, and $\kappa^{(pp)}$ and $\kappa^{(mp)}$ are the intra- and inter-modal averaging factors, given by $\frac{4}{3}\frac{2M}{2M-1}$ and $\frac{4}{3}$, respectively [Fe16]. $A_{eff}^{(pppp)}$ and $A_{eff}^{(mmpp)}$ are the intra-modal and inter-modal effective area, respectively, as defined in [Ra16.1].

For the sake of clarity, in the rest of the presented equations, superscripts define spatial mode indexing and subscripts define wavelength indexing. The interaction of three wavelengths λ_1, λ_2, and λ_3, propagating in different spatial modes denoted respectively as m, n, and p, generates a nonlinear signal $E_{\lambda_4}^{(r)}$at a frequency $\omega_{\lambda_4} = \omega_{\lambda_1} + \omega_{\lambda_2} - \omega_{\lambda_3}$ in spatial mode r as a consequence of FWM [El13], [Fe16]:

$$E_{\lambda_4}^{(r)} = \tilde{\gamma}^{(mnpr)}E_{\lambda_1}^{(m)}E_{\lambda_2}^{(n)}E_{\lambda_3}^{(p)}\frac{1-e^{-\alpha L}e^{-j\Delta\beta_{\lambda_1\lambda_2\lambda_3\lambda_4}^{(mnpr)}L}}{j\Delta\beta_{\lambda_1\lambda_2\lambda_3\lambda_4}^{(mnpr)}L+\alpha}e^{-\alpha/2\,L}e^{-j\beta}ds^L \quad (9\text{-}4)$$

where L is the span length, $\tilde{\gamma}^{(mnpr)}$ is the nonlinear coefficient for the interaction of spatial modes m, n, p, r, and $\Delta\beta_{\lambda_1\lambda_2\lambda_3\lambda_4}^{(mnpr)} = \beta_{\lambda_1}^{(m)} + \beta_{\lambda_2}^{(n)} - \beta_{\lambda_3}^{(p)} - \beta_{\lambda_4}^{(r)}$ is the phase mismatch between wavelengths λ_1, λ_2, λ_3, and λ_4, with $\beta_{\lambda_1}^{(m)}$, $\beta_{\lambda_2}^{(n)}$, $\beta_{\lambda_3}^{(p)}$, and $\beta_{\lambda_4}^{(r)}$

being the propagation constant of spatial mode m, n, p, and r at the angular frequency ω_{λ_1}, ω_{λ_2}, ω_{λ_3}, and ω_{λ_4}, respectively.

The total nonlinear power generated by the effect of FWM for a Sp-Ch with total bandwidth B can be obtained by integrating the product of the electrical field in Eq. (8-4) and the signal Power Spectral Density (PSD) in each spatial mode over frequencies ω_{λ_1}, ω_{λ_2}, and ω_{λ_3}. The closed-form of this integral is derived in [Hu16] under the assumption that the spectrum signal is rectangular as in the case of Nyquist WDM. The derivation of the closed-form has led to the definition of the following FWM efficiency parameter [El13], [Fe16]

$$\eta^{(mnpr)} = \frac{[\tilde{\gamma}^{(mnpr)}]^2}{2\pi\alpha\left|\beta_2\right|}\left[\ln\left(\frac{B^2 + 2B\Delta f^{(mnpr)}}{2f_w^2}\right) + \bar{S}\ln\left(\frac{\bar{S}B^2 - 2B\Delta f^{(mnpr)}}{2f_w^2}\right)\right] \tag{9-5}$$

The term f_w is the walk-off bandwidth indicating the frequency range of the effectiveness of FWM nonlinearity in the presence of the dispersion defined as:

$$f_w = \sqrt{\frac{\alpha}{4\pi^2\left|\beta_2\right|}} \tag{9-6}$$

The terms $\bar{S}$ and $\Delta f^{(abcd)}$ are respectively defined as:

$$\bar{S} = sign(B - 2\Delta f^{(mnpr)}) \tag{9-7}$$

$$\Delta f^{(mnpr)} = \frac{(\beta_1^{(m)} + \beta_1^{(n)} - \beta_1^{(p)} - \beta_1^{(r)})}{2\pi\beta_2} \tag{9-8}$$

The total nonlinear PSD in spatial mode r in an M-mode FMF is then given by:

$$G_{NLI}^{(r)} = \sum_{m=1}^{M}\sum_{n=1}^{M}\sum_{p=1}^{M}\eta^{(mnpr)}G^{(m)}G^{(n)}G^{(p)} \tag{9-9}$$

where $G^{(m)}$, $G^{(n)}$, and $G^{(p)}$ are, respectively, the PSDs of spatial modes m, n, and p. Only interactions between pairs of spatial modes are considered (i.e., $m = r, n = p$ or $n = r, m = p$) since the other interactions give rise to unstable FWM processes that can be neglected [Ra16.1]. Under this assumption, Eq. (8-9) becomes:

$$\begin{aligned} G_{NLI}^{(r)} &= G^{(r)}\left\{\eta^{(rrrr)}[G^{(r)}]^2 + 2\sum_{n=1,n\neq r}^{M}(\eta^{(rnnr)} + \eta^{(nrrn)})[G^{(r)}]^2\right\} \\ &= G^{(r)}\left\{\eta^{(rrrr)}[G^{(r)}]^2 + \sum_{n=1,n\neq r}^{M}(4\eta^{(rnnr)}[G^{(r)}]^2)\right\} \end{aligned} \tag{9-10}$$

The first term represents the intra-modal nonlinear PSD, while the second term is the total inter-modal nonlinear PSD given by the contribution of the interacting spatial modes. The factor of 2 in the first row of the equation is because two polarizations have been considered for each mode. The second row is obtained considering that $\eta^{(rnnr)} = \eta^{(nrrn)}$ since we only have interactions between pairs of spatial modes.

The total nonlinear power $P_{NLI}^{(r)}$ in spatial mode r and the ASE noise power P_{ASE} in the amplifiers are the main sources of degradation of the QoT decreasing the OSNR. Considering an FMF link composed of N_s spans of identical lengths, we obtain:

$$P_{NLI}^{(r)} \cong N_s G_{NLI}^{(r)} B_{ref} \tag{9-11}$$

and

$$P_{ASE} \cong N_s F_n h\nu G B_{ref} = N_s G_{ASE} B_{ref} \tag{9-12}$$

where F_n is the noise figure of the amplifiers, $h\nu$ is the photon energy, B_{ref} is the reference noise bandwidth, G is the amplifier gain, and G_{ASE} is the noise PSD in the amplifiers, assumed the same for all spatial modes in our model.

Then, the $OSNR_{FMF}^{(r)}$ for spatial mode r in an FMF is [Po14]:

$$OSNR_{FMF}^{(r)} = \frac{P_{ch}}{N_s (G_{ASE} + G_{NLI}^{(r)} B_{ref})} \tag{9-13}$$

where P_{ch} is the input channel power.

Eq. (8-13) can also describe BuSMFs, considering P_{NLI} only accounting for intra-modal nonlinearities.

Current nonlinear models [El13], [Fe16], [Ra16.2] assume the same input power for all spatial modes, ideal span-by-span optical amplification that perfectly compensates for the loss of the previous span, and spans with equal lengths and same loss coefficient for all spatial modes. In a realistic scenario, however, the overall loss that each spatial mode path experiences are different due to MDL in the fiber and the multiplexing and switching components, as well as the Mode-Dependent Gain (MDG) of the FM-EDFAs [Ba11]. An attempt to account for the impact of MDL on a networking study has been presented in [Hu16] based on a model for strongly-coupled FMFs [Ho12] without considering nonlinearities. However, in most practical cases, one will be interested in FMF systems with up to 6 spatial modes, which are mainly considered to operate in the weakly-coupled regime (as in the nonlinear models presented in [El13], [Fe16], [Ra16.2]). As of today, no analytical model including both the impact of nonlinearities and MDL/MDG have been presented so far due to the increased complexity (and

therefore impracticability) of such a model. As a result, the following QoT estimator only focuses on current analytical nonlinear models that exclude the impact of MDL/MDG on the system performance [El13], [Fe16].

9.2.2 An Analytical Model for MCFs

The FWM effect due to signals propagating in different cores is defined in a similar way to the case of FMFs. Indeed, considering single mode cores and three wavelengths λ_1, λ_2, and λ_3 propagating in cores m, n, and p, respectively, a signal is generated at frequency $\omega_{\lambda_4} = \omega_{\lambda_1} + \omega_{\lambda_2} - \omega_{\lambda_3}$ in core r, with a FWM efficiency given by

$$\eta^{(mnpr)} = \frac{[\tilde{\gamma}^{(mnpr)}]^2}{\pi\alpha\left|\beta_2\right|} \ln\left(\frac{B^2}{2 f_w^2}\right) \tag{9-14}$$

where f_w is the walk-off bandwidth, $\tilde{\gamma}^{2\,(mnpr)}$ is the nonlinear coefficient, α is the attenuation coefficient, β_2 is the GVD and B is the overall bandwidth of the channel. Therefore, as is the case for FMFs, the total nonlinear PSD in core r is:

$$G_{NLI}^{(r)} = \sum_{m=1}^{C}\sum_{n=1}^{C}\sum_{p=1}^{C} \eta^{(mnpr)} G^{(m)} G^{(n)} G^{(p)} \tag{9-15}$$

where $G^{(m)}$, $G^{(n)}$, and $G^{(p)}$ are the power spectral density in core m, n, and p, and C is the total number of cores. The above equation can be derived easily by setting the velocity-matched offset $\Delta f^{(mnpr)}$ to zero in Eq. (8-8). This is a valid assumption for MCFs with homogeneous cores with equal mode velocities.

The total nonlinear power $P_{NLI}^{(r)}$ and ASE noise power P_{ASE} in the amplifiers are also given by Eq. (8-11) and (8-12), as in the case of FMFs.

When it comes to the estimation of the impact on one core of its surrounding cores in an MCF, the fabrication parameters need to be taken into consideration. Depending on the distance between cores (and, consequently, the number of cores) and other parameters, such as the refractive index profile of each core, MCFs can have strongly-coupled or weakly-coupled cores. In the former case, a MIMO DSP is required, which will ideally correct all effects of linear Crosstalk (XT) between cores. Conversely, in the latter case, one will lean toward doing away with the MIMO DSP to bring down dispensable expenses, and linear XT will necessarily have to be considered for the evaluation of the OSNR degradation. Therefore, unlike in the FMF case, the OSNR formula will have to factor in the following additional term [Ch10], [Ha12]:

$$P_{XT}^{(r)} \cong N_s \mu_{XT}^{(r)} P_{ch} \tag{9-16}$$

where N_s is the number of spans and $\mu_{XT}^{(r)}$ is the XT parameter given by

$$\mu_{XT}^{(r)} \cong \sum_{n=1, n\neq r}^{C} \eta_{coupl}^{(nr)} L_{span} \tag{9-17}$$

L_{span} being the span length and $\eta_{coupl}^{(nr)}$ being the power coupling coefficient between cores n and r [Ha12], [Ko12], [Ye14].

In the case of homogeneous MCFs, the previous equation can be simplified as expressed by the following formula:

$$\mu_{XT}^{(r)} \cong N_c \eta_{coupl} L_{span} \tag{9-18}$$

where N_c is the number of cores adjacent to core r.

Taking all these terms into account, the OSNR for core r in a homogenous weakly-coupled MCF can be written as

$$OSNR_{MCF}^{(r)} = \frac{P_{ch}}{N_s (G_{ASE} + G_{NLI}^{(r)}) + P_{XT}^{(r)}} \tag{9-19}$$

9.2.3 Simulation Environment

The QoT estimator is designed to receive as input some parameters providing information about the transceiver model and the transmission channel. The output of the tool produces fundamental metrics in establishing impairment-aware optical connections.

The estimation mechanism requires input parameters to describe the transmitting system, such as input power per channel (Pch), number of co-propagating spectral channels (Nch), number of links ($Nlinks$), number of spans in a link ($Nspans$), span length ($Lspan$), symbol rate (R_s), channel spacing (Δf), and modulation format ($ModFormat$). Additional parameters are needed according to the model to be used for the simulations (FMF, MCF or BuSMF), and also to specify the characteristics of the fiber employed as the transmission channel.

For the FMF case, the QoT estimator requires closed-form equations (8-11) and (8-12) to calculate the total nonlinear interference noise power $P_{NLI}^{(r)}$ and P_{ASE} for each fiber spatial mode d. For the MCF case, the QoT estimator requires equations (8-11), (8-12) and (8-16) to calculate $P_{NLI}^{(r)}$, P_{ASE} and the linear crosstalk power $P_{XT}^{(r)}$ for fiber core r. For the BuSMFs case, the same equations apply as for the MCF case, with $P_{XT}^{(r)} = 0$ and $P_{NLI}^{(r)}$ only accounting for intra-modal contributions. The Signal-to-

Noise Ratio (SNR) corresponding to the $OSNR_{FMF}^{(r)}$ or $OSNR_{MCF}^{(r)}$ value for spatial mode or core d, obtained as reported in Eq. (8-13) and (8-19), respectively, can be calculated as:

$$OSNR_{ASE,NLI} = \frac{OSNR_{FMF\,(\text{or } MCF)}^{(r)} \times B_{ref}}{N_p \times B_{rx}} \tag{9-20}$$

where B_{ref} is the reference noise bandwidth, N_p is the number of polarizations and B_{rx} is the receiver filter bandwidth. The total SNR, in turn, is given by:

$$SNR = \frac{1}{\frac{1}{SNR_{XT}} + \frac{1}{SNR_{ASE,NLI}}} \tag{9-21}$$

where SNR_{XI} is the SNR value accounting for inter-channel XT (between spectral channels within the same spatial mode) and inter-symbol interference in back-to-back measurements for different modulation formats and spectral widths of the signal [Ho12].

BER values can be calculated from SNR_{XI} for different modulation formats, according to the following formulas [Ca12]:

- DP-BPSK $$BER = \frac{1}{2} erfc\left(\sqrt{SNR}\right) \tag{9-22}$$
- DP-QPSK $$BER = \frac{1}{2} erfc\left(\sqrt{\frac{1}{2} SNR}\right) \tag{9-23}$$
- DP-8QAM $$BER = \frac{2}{3} erfc\left(\sqrt{\frac{3}{14} SNR}\right) \tag{9-24}$$
- DP-16QAM $$BER = \frac{3}{8} erfc\left(\sqrt{\frac{1}{10} SNR}\right) \tag{9-25}$$

where $erfc(x)$ is the complementary error function $erfc\,(x) = 2/\sqrt{\pi} \int_x^{\infty} \exp(-t^2)\, dt$.

The Q-factor can be computed from the BER as follows:

$$Q = 20 \log_{10}\left(\sqrt{2} erfc^{-1}(2 \times \text{BER})\right) \tag{9-26}$$

where $erfc^{-1}(x)$ is the inverse of the complementary error function.

9.2.4 Transmission Model

The transceiver model considered for the simulations is a WDM Sp-Ch composed of N_{ch} co-propagating spectral channels with symbol rate R_s and subcarrier spacing of Δf. The simulated transceivers are characterized by four different modulation formats: DP-BPSK, DP-QPSK, DP-8QAM, and DP-16QAM. The transmission channel is represented by a multi-span FMF or MCF spans. After each fiber span, an EDFA with a noise figure (F_n) equal to 5 dB completely compensates the attenuation of the preceding span.

In the case of FMF, the reference fiber considered for the simulations is an FMF supporting four non-degenerate LP spatial modes, i.e., a six-mode FMF with spatial modes LP_{01}, LP_{11a}/LP_{11b}, LP_{02}, and LP_{21a}/LP_{21b}. For this fiber, $M = 6$ and the intra-modal ($\kappa^{(pp)}$) and inter-modal ($\kappa^{(mp)}$) averaging factors in Eq. (8-2) and (8-3) are equal to $\frac{16}{11}$ and $\frac{4}{3}$, respectively. Only the interaction between pairs of spatial modes is considered in the model presented here. Therefore, the only inter-modal effective area ($A_{eff}^{(mp)}$) values required for the estimation of equations (8-2) and (8-3) are presented in the matrix in Table 9-1. Likewise, for the calculation of Eq. (8-8), Table 9-2 shows the relative differential mode delay for all six spatial modes with respect to LP01.

Table 9-1 Effective area [μm²] for a 6-mode FMF

	LP01	LP02	LP11a	LP11b	LP21a	LP21b
LP01	146	291	291	291	583	583
LP02	291	290	578	578	578	578
LP11a	291	578	193	579	386	386
LP11b	291	578	579	193	386	386
LP21a	583	578	386	386	257	772
LP21b	583	578	386	386	772	257

Table 9-2 Differential mode delay [ps/km] between pairs of modes of a 6-mode FMF

	LP01	LP02	LP11a	LP11b	LP21a	LP21b
LP01	0	8	13	14	17	18

Finally, we assume that all the LP modes are affected by the same amount of CD, equal to 20 ps/nm/km, and the attenuation parameter and nonlinear index equal $\alpha = 0.2\,dB/km$ and $n_2 = 2.6 * 10^{-20}\,m^2/W$, respectively. These values are summarized in Table 9-3.

For the MCF simulations, we assume a seven-core MCF with identical hexagonally-arranged pure-silica cores [Ha12]. Since the $P_{XT}^{(d)}$ in each core depends on the contribution of adjacent cores, as shown in Eq. (18), two cases can be distinguished for this fiber configuration:

- XT in the center core, representing the worst-case scenario since the center core has six adjacent cores ($N_c = 6$);
- XT in the outer cores, affeted only by the center core and the two adjacent outer cores ($N_c = 3$).

We consider the values for the statistical mean of the XT $\mu_{XT}^{(d)}$ at $\lambda = 1550\,nm$ extracted from Table 2 in Ref [Ha12]. $\mu_{XT}^{(d)}$ is used to conveniently compute the power coupling coefficient $\eta_{coupl,bd}$. The effective area is fixed to $75\,\mu m^2$, CD to 16.7 ps/nm/km, and nonlinear index (n_2) to $2.3 * 10^{-20}\,m^2/W$. The attenuation parameter is again set to $\alpha = 0.2\,dB/km$. These parameters are presented in Table 9-4.

Table 9-3 FMF parameters for simulations

α[dB/km]	CD[ps/nm/km]	n_2[m²/W]	k_{pp}	kmp
0.2	20	2.6×10^{-20}	16/11	4/3

Table 9-4 MCF parameters for simulations

α[dB/km]	CD[ps/nm/km]	n_2[m²/W]	Aeff[μm²]	kmp
0.2	20	2.6×10^{-20}	16/11	4/3

9.3 **Simulations Results**

The SDM QoT estimator allowed to quantitatively compare the performance of FMFs, MCFs, and BuSMFs and therefore to evaluate the impact of the physical impairments. The analysis here extends the results presented in [Ma17.2]. A 6-mode FMF, a7-core weakly-coupled MCFs, and a bundle of 6 SMFs were considered, and their performance was compared in terms of maximum optical reach. The fiber parameters used are those reported in Table 9-3 and Table 9-4.

In order to evaluate the penalty introduced by the use of more complex modulation formats, we consider the four modulation formats described in the previous section. To determine the maximum reach that a Sp-Ch can achieve, we assumed a multi-span system. The transmission reach estimation depends on the considered span length, since the value of the BER is calculated span by span, and then the obtained value is compared with the target BER. In addition to span length, the number of co-propagating channels and the channel spacing are two parameters

affecting the transmission reach. We performed a set of simulations determining the impact of these parameters on the achievable reach.

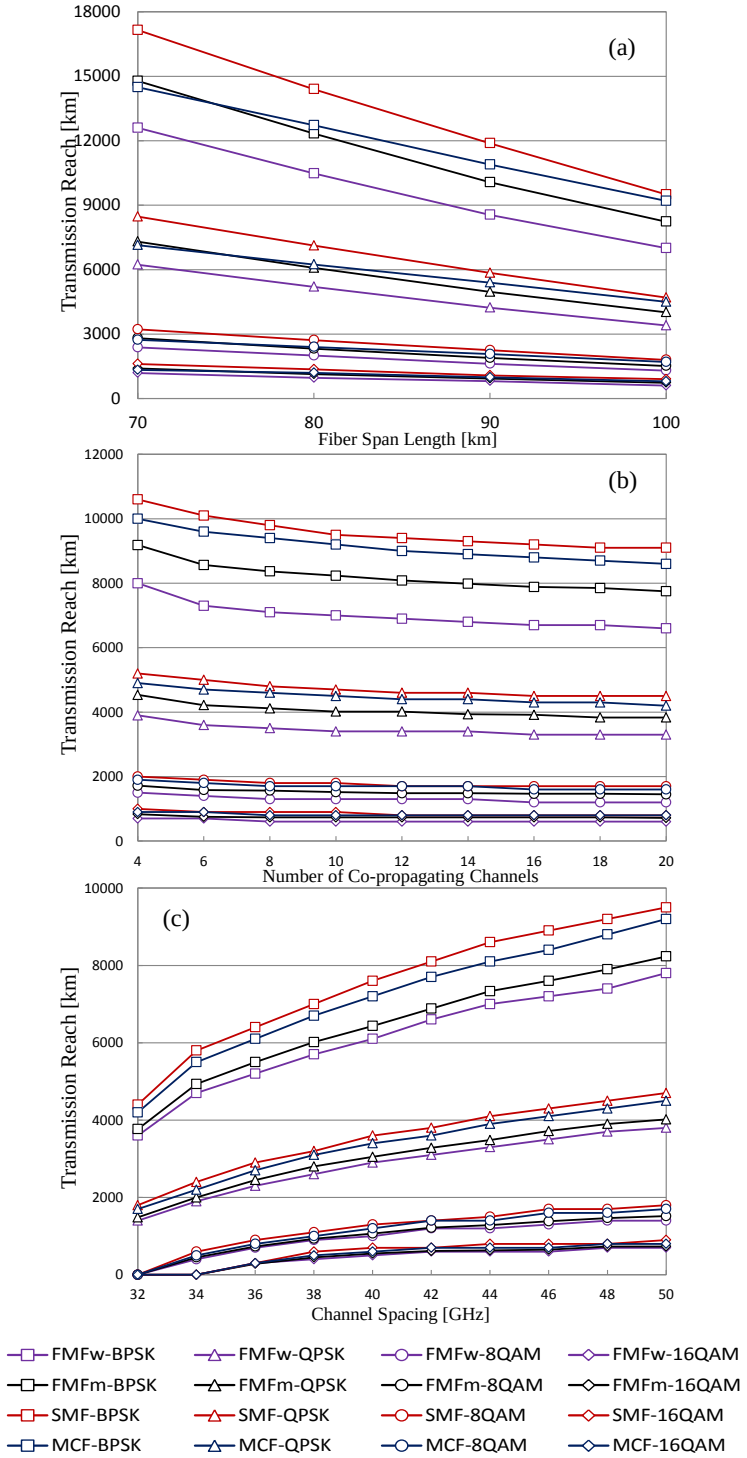


Fig. 9-1. The maximum reach as a function of (a) fiber span length, (b) the number of co-propagating channels, and (c) channel spacing

The results are shown in Fig. 9-1. For the MCFs the depicted values are referred only to the center core. A mean distance value over the six modes and the worst-case value are considered for the FMFs. In Fig. 9-1(a) are reported the maximum reach values obtained varying the span length. The simulations were performed fixing the symbol rate to 32 GBaud, the channel spacing to 50 GHz and the number of co-propagating channels to 10. Each value is obtained considering the optimal power for that case.

Fig. 9-1(a) shows how the reach decreases when the length of the span increases. Note that, in the MCF case, for higher powers (when the amplification span length is shorter) the linear crosstalk is higher. So it shows a poor performance compared to BuSMFs for short spans, while it converges to the performance of BuSMFs for long spans. The impact of the fiber span length increasing on the transmission reach is stronger when less complex modulation formats as BPSK or QPSK are employed. The picture also shows the maximum reach values that can be achieved in each scenario with a given span length.

Fig. 9-1(b) describes the variation of the transmission reach when the number of co-propagating channels composing the Sp-Ch is changed. 100 km span length is considered for the results shown in Fig. 9-1(b) and (c). The considered channels are 50 GHz spaced. From the graph, it is visible that high-level modulation formats incur a higher OSNR penalty, resulting in shorter transmission reach, but are more robust to the increase of the number of co-propagating channels since no significant decrease in the optical reach was measured.

Finally, Fig. 9-1(c) shows the effect of channel spacing variation on the maximum optical reach for the optimum launch power. An increase of the transparent reach is accomplished by increasing the spacing between subcarriers in the Sp-Ch, even if this results in higher bandwidth occupancy for the Sp-Ch with a consequent decrease in the spectral efficiency.

9.4 **End-to-End QoT Estimator**

From a practical point of view, there are multiple active connections in a network, which results in links accommodating a different number of channels from different end-to-end connections. Therefore, for a given lightpath the signal related characteristics are not the same for each link. The QoT estimator module is designed to provide a separate estimation for each point-to-point link in between nodes. Then, the overall performance of the lightpath is estimated by concatenating the degradation effects from each link in the OSNR values according to [Po14].

QoT estimator works in cooperation with other modules of the offline planning tool and in particular with the RSMSA module to minimize the spectrum utilization and spectrum fragmentation, while maintaining the required maximum BER. The QoT-aware offline planning tool thus accounts for physical layer impairments of

different types of SDM fibers and the extra constraints imposed by SDM switching technologies.

9.5 Network-wide Performance Evaluation

Selection of a proper transmission medium is one of the key decisions to be made for the development of SDM networks. In principle, the proper option is the one, which not only offers good performance but also imposes a reasonable level of component complexity and deployment cost. This is because the complexity of the components (transceiver, switches, amplifiers, etc.) realizing a particular flavour of SDM network is highly fiber-dependent. Studying the complexity of components required for SDM networks based on different types of fibers is out of the scope of this work. However, it is of utmost importance to investigate the performance of SDM networks based on the different fiber options pointed out in this work to showcase the benefits of the developed QoT tool for SDM-based optical network planning practices.

In order to carry out the studies, we consider the topology of Telefónica Spain national network, and it's traffic matrix as described in Section 8.2.2. In our analysis, we assume two versions of the topology; one with the original link lengths, with average link length equal to 135 km, and one where we double the link lengths to evaluate the impact of the geographical size of the network on the performance of SDM networks utilizing alternative transmission media. We should also highlight that even though there are limitations with the adopted physical-layer model since it just considers the operation in the weak coupling regime, the distances we are considering in our analysis are indeed possible with proper SDM system configurations, as it has been demonstrated in recent works related with transmission experiments using FMFs for long-haul applications [Ra18.2].

As mentioned in Section 3.5, the choice of a transmission medium is a key factor in determining the required architecture of the optical switches in an SDM-based network. BuSMFs and weakly-coupled MCFs are compatible with all SDM switching paradigms described in the previous chapter, while it is necessary to consider J-Sw when FMFs are in place, since all the spatial modes of FMFs have to be switched together and received by the same receiver allowing the MIMO-based DSP to correctly retrieve the inter-mixed information across different spatial modes [Ma15.1].

Therefore, we choose J-Sw as the switching technology for the studies presented in this chapter, since it is compatible with all SDM fiber options. Note that selecting this option allows us to focus solely on the performance of transmission media, which is the goal of this study, regardless of the flexibility/constraints imposed by switching schemes, whose impact is thoroughly investigated before and presented in the previous chapter.

Regarding the transmission technology, spatial Sp-Ch allocation scheme is used, which is also compatible with the J-Sw scheme. For the networking study, we consider a 6-mode FMF, a bundle of 6 SMFs, and a 7-core MCF in which we only use the six outer cores to have the same amount of available capacity per link regardless of the fiber type. For the FMF case, two baselines for estimating the QoT are assumed: 1) based on the average (mean) of QoT values of the 6 modes, and 2) based on the worst QoT values, i.e., LP21a/b, which are labelled as "FMFm" and "FMFw" in the figures, respectively. We consider two metrics for comparison purposes; the average spectrum utilization of the network and the corresponding number of required transceivers operating over different modulation formats.

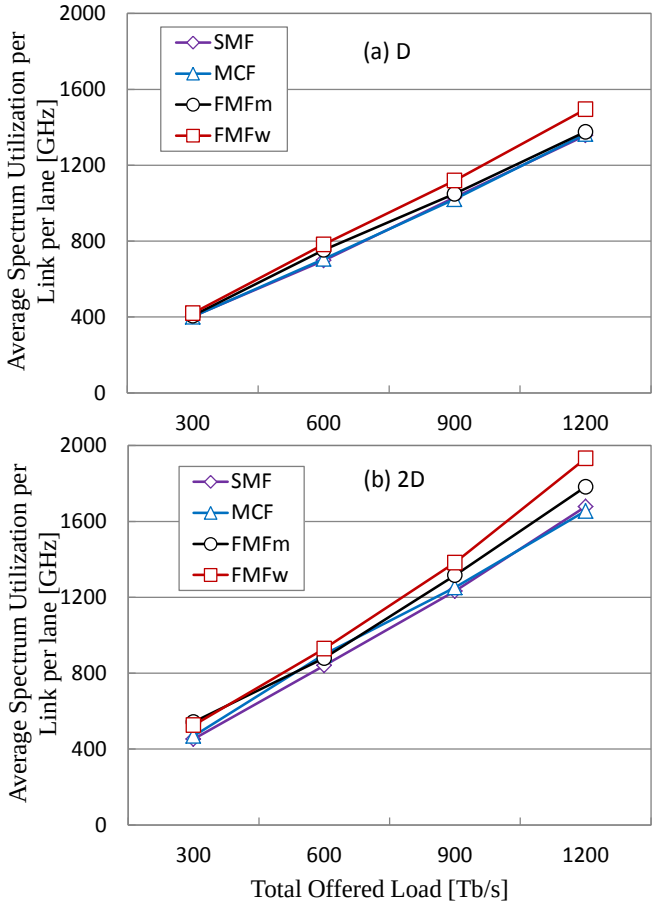


Fig. 9-2. The average spectrum utilization for (a) the actual sized topology (D) and (b) the doubled sized topology (2D)

Fig. 9-2 shows the networking level performance in terms of the average spectrum utilization for the two different sizes of topology under study; (a) the actual sized topology (D) and (b) the doubled sized topology (2D). Note that, in this context, the performance improvement is achieved when the average spectrum utilization decreases. The performance of networks based on BuSMFs and MCFs is quite

similar, even though a slight degradation in the performance of networks based on MCFs is observed in the case of the large-scale topology (Fig. 9-2).

This degradation is more evident while looking at the number of required TRxs operating over different modulation formats plotted in Fig. 9-3. Considering the large-scale topology (Fig. 9-3 (b)), the case utilizing MCFs uses a higher number of less spectrally efficient TRxs (8QAM compared to 16QAM) compared to the one based on BuSMFs resulting in a slightly higher number of TRxs for the MCF case. This difference is also shown in Fig. 9-4 and Fig. 9-5.

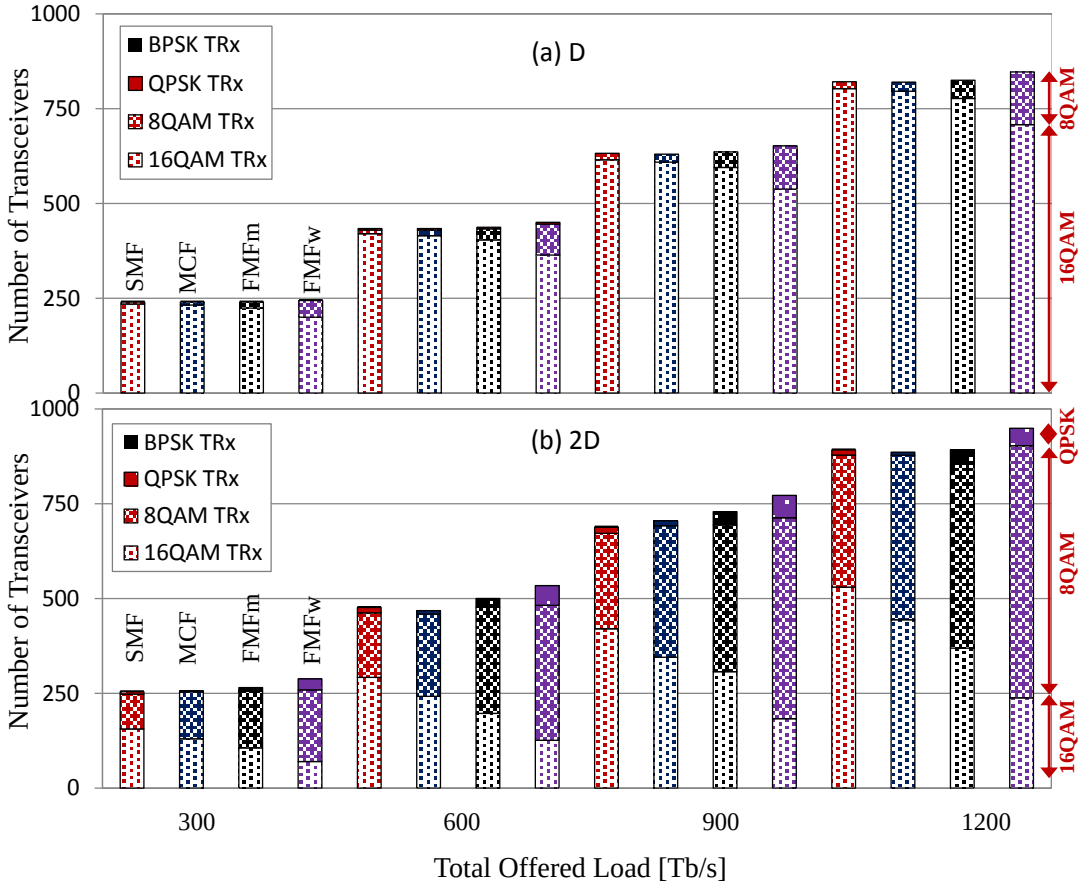


Fig. 9-3. The number of required transceivers for (a) the actual sized topology (D) and (b) the doubled sized topology (2D).

While considering networks utilizing FMFs based on the worst QoT value (FMFw), a significant difference between its performance with the cases of SMFs and MCFs is observed. This performance difference becomes more pronounced for the larger network compared to the smaller one. This performance difference is well illustrated in Fig. 9-3 and Fig. 9-5. The FMFw case not only requires many more TRxs, regardless of the topology size, but also necessitates the use of much lesser-spectrally efficient TRxs compared to the other transmission media. In the large-

scale topology, it requires 90 more TRxs in total compared to the BuSMFs-based network.

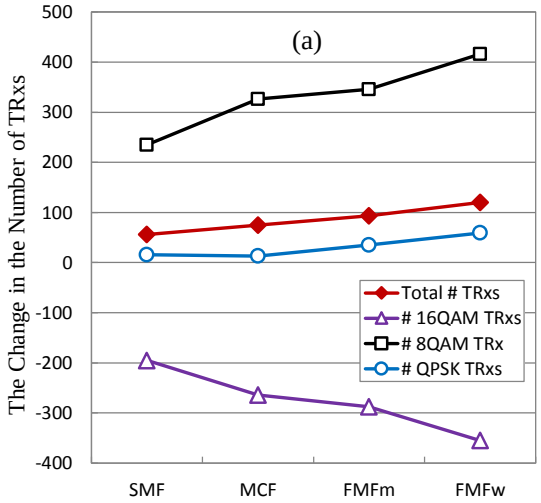


Fig. 9-4. The change in the number of TRxs while considering the large-scale topology compared to the small one for every fiber type independently. The total offered load is 900 Tb/s.

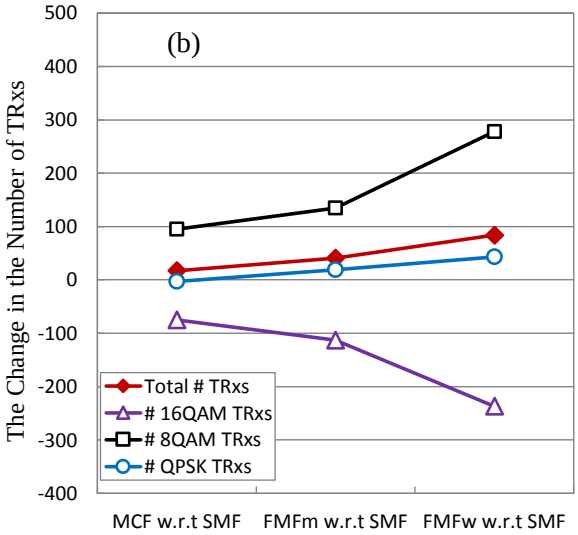


Fig. 9-5. The change in the number of TRxs considering MCFs/FMFs compared to SMFs for the large-scale topology. Every point is obtained by subtracting the number of required TRxs for BuSMFs case from the target case.

In addition, out of the total number of in-operation TRxs, only ~24% can operate over 16QAM format, while this approaches 60% for the case of BuSMFs. In other words, by using FMFs instead of SMFs, the number of in-operation TRxs over 16QAM decreases sharply and instead more TRxs with lower spectral efficiency are launched resulting in a higher amount of spectrum utilization in the network.

Considering spatially integrated Sp-Ch TRxs, it is expected that a single common FEC code is used for all spatial channels, enabling the signals traversing the FMF-based network to travel as far as the average of the reach values of all six modes. Therefore, to show the performance benefits of such option enabled by integrated TRxs, we consider this case in our analysis (labelled as FMFm). Under such circumstances, it is observed that the performance of networks based on FMFs can be enhanced significantly and approaches the performance of the cases of BuSMFs and MCFs for small-scaled topologies. Additionally, for the large topology, it can greatly improve the performance of networks based on FMFs. As shown in Fig. 9-3, it requires deploying fewer TRxs compared to FMFw case for the same amount of offered load to the network.

In addition, as shown in Fig. 9-5, the number of TRxs operating over more spectrally efficient modulation formats increases compared to FMFw. However, it yet performs quite badly compared to BuSMFs and weakly-coupled MCFs for large-scale topologies. Therefore, even though the introduction of a common FEC code may enhance the performance of FMF-based SDM networks and make them favorable for small-scale topologies, they lack enough potential to be considered for large-scale topologies.

Regardless of their performance, the DSP complexity and its corresponding power consumption may be a showstopper for FMF-based SDM networks. On the other hand, other components used for networks based on FMFs like integrated amplifiers (e.g. FM-EDFAs) and switching nodes can reduce power consumption. Note that, while the DSP complexity in the case of BuSMFs and weakly-coupled MCFs is similar to that of conventional SMF-based systems, it is higher for the FMF cases because of the severe effects of DGD, which in turn results in higher cost and power consumption of the transceivers.

9.6 **Concluding Remarks**

In this chapter, we have compared the performance of SDM networks utilizing different types of fibers. More specifically, we have investigated the impact of physical layer impairments on the performance of SDM networks deployed over different types of transmission media. In this regard, we first presented a quality of transmission estimator using accurate, yet low computationally complex, analytical models for SDM networks.

Based on this estimator, the performance of SDM networks using different transmission media has been compared. In the analysis, we considered two different sizes of Telefonica Spain topology. Considering small-scale topologies, FMF-based SDM networks show slightly worse performance compared to the case of BuSMFs and weakly-coupled MCFs. However, for large-scale topologies, SDM networks based on BuSMFs and MCFs show a significantly better performance,

mainly in terms of the amount of utilized spectral and spatial resources, which allows the network to support more traffic compared to FMF-based realizations.

It can be concluded that FMFs can be an option to be considered for small-scale networks, even though it requires a more complex MIMO-DSP, which translates into higher implementation cost and power consumption. However, for large-scale networks, BuSMFS and MCFs are shown to be the most favourable options for large-scale topologies. Additionally, the utilization of the already-installed SMFs is likely to be the most cost-effective option for near-term realizations of SDM networks. However, in practice, the splicing of BuSMFs with a large number of fibers and their alignment with the TRx is a challenge since it makes the interfacing more difficult and costlier compared to the FMF case. MCFs also pose the same problem. On the other hand, MCFs and FMFs offer reduced-size cables with higher density and lower weight then BuSMFs.

Chapter 10

Node Design and Power Consumption Analysis

In this chapter, we aim at complementing the studies presented in the last two chapters by analyzing the impact of extra complexity introduced due to spatially integrating the switching nodes and transmission media.

We have already shown that the SDM switching strategies have a significant impact on the performance of SDM networks. Even though some of them show superior performance, they impose high levels of complexity for the realization of the corresponding SDM ROADMs. Therefore, as one of the contributions of this chapter, we explore the impact of such complexity on the realization of SDM ROADMs and find the most appropriate designs, which limit such complexity and reduce their implementation cost.

In addition, we discussed that SDM networks based on integrated fibers (e.g. FMFs) require sophisticated MIMO-DSP algorithms. In our analysis, we considered that proper MIMO-DSP algorithms can perfectly correct all the linear impairments. However, the MIMO-DSP algorithms also impose a significant level of complexity to the system. The higher the spatial integration of the spatial modes in a fiber, the higher is the amount of linear impairments. Therefore, as another contribution of this chapter, we investigate the impact of the increased complexity on the system performance as well as power consumption of the TRxs.

10.1 Comparison of ROADM Architectures for SDM Networks

Commercially speaking, CDC ROADMs are becoming a reality for conventional flexible WDM networks. The existing architectures can be upgraded to address the challenges introduced by SDM, which will arguably be the next step in the

evolution of optical networks. These ROADMs can be realized following one of the three SDM switching strategies described before: Ind-Sw, FrJ-Sw, and J-Sw, in order of decreasing flexibility and increasing hardware efficiency.

Fig. 10-2 and Fig. 10-2 show two route and select ROADM architectures (with nodal degree D=3) enabling Ind-Sw and J-Sw, respectively, in an SDM network with two spatial dimensions (S=2). In Fig. 10-2, we show all internal connectivity from the 'East' ingress WSS. LCs are supported if all solid, dashed and dotted connections exist. Ind-Sw without LCs requires solid and dashed lines. In Fig. 10-2, each spatial dimension is routed independently. Solid lines represent locally added or dropped traffic, dashed lines depict pass-through connections between WSS belonging to the first (grey) or second (blue) dimensions, and red dotted lines account for the possibility of having connections between different spatial dimensions (referred to as LC in [Ma15.1]). Conversely, in Fig. 10-2, all spatial dimensions are routed together. An intermediate solution is FrJ-Sw, whereby a group of G spatial dimensions (G being a divisor of S) is jointly routed (not shown).

We propose several colorless directionless and CDC ROADM architectures enabling Ind-Sw, FrJ-Sw and J-Sw of spatial Sp-Chs and we present a cost model for comparing the proposed architectures. We find that a CDC ROADM design, which maximizes the number of A/D ports, and yet keeps the port count of pass-through WSSs low, is the most cost-effective solution.

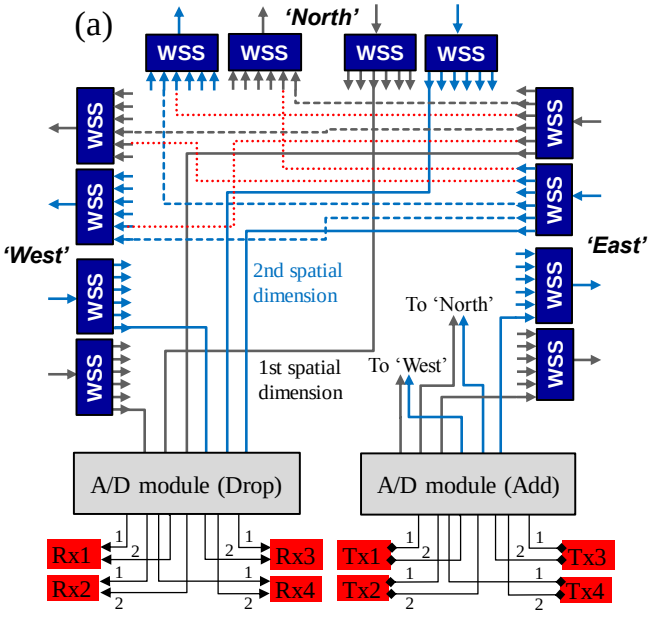


Fig. 10-1. Route and select ROADM architectures for Ind-Sw

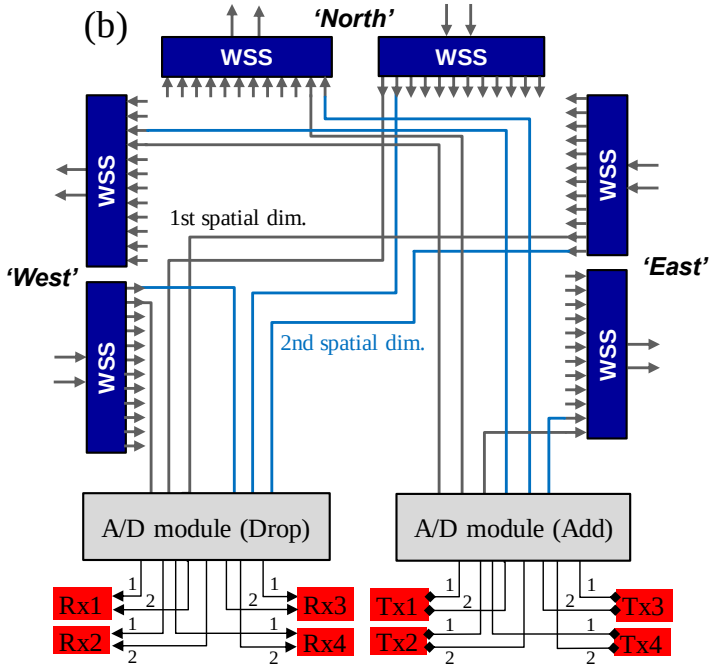


Fig. 10-2. Route and select ROADM architectures for J-Sw with S=2.

10.1.1 ROADM Architectures and Cost Analysis

We consider the ROADM architectures illustrated in Fig. 10-2 and Fig. 10-2, as well as the case (not shown) of FrJ-Sw, where each group of G spatial dimensions is jointly routed by a $G\times(I\times O)$ WSS. LCs between groups can also be supported in the case of FrJ-Sw with the appropriate connectivity and port count. The required number of WSSs and their port count is shown in Table 10-1 as a function of G (thus providing a common formula valid for all cases) for routing options with/without LC. The evaluation of the ROADM cost (C_{ROADM}) considers the cost of the WSSs (C_{WSS}) and an A/D module ($C_{\text{A/D}}$) composed of K elements, such that K minimizes the overall cost.

Table 10-1 Number of elements and total cost of the ROADM architectures (Fig. 10-2)

	Pass-through WSS		Total ROADM cost (C_{ROADM})
	#	WSS port count	
Without LC	$2 \cdot D \cdot S/G$	G×[1×(D+K-1)]	$\min\{C_{A/D}+2D\cdot(S/G)\cdot C_{WSS(G\times[1\times(D+K-1)])}\}$
With LC		G×{1×[(S/G)×(D-1)+K)]}	$\min\{C_{A/D}+2D\cdot(S/G)\cdot C_{WSS(G\times\{1\times[(S/G)\times(D-1)+K)]\})}\}$

For C_{WSS} we use as a baseline the cost of commercial LCoS-based 1×9 WSSs (cost = 1). Node architectures requiring WSSs with port counts 1×5, 1×9, and 1×20, all commercially available, have a cost per WSS of 0.63, 1 and 1.58, respectively. The

cost of WSSs with port counts 1×40, 1×80 and 1×160 was extrapolated to be 2.50, 3.95, and 6.25, respectively, according to 58% premium for a fiber port doubling.

The A/D modules are implemented as shown in Fig. 10-3 for CD operation and Fig. 10-4 - Fig. 10-5 for CDC operation (only the drop module is shown). Architectures for Colorless Contention-less (CC) operation were investigated in [Ma15.1]. In Fig. 10-3 - Fig. 10-5, T is the number of TRxs connected to the A/D ports.

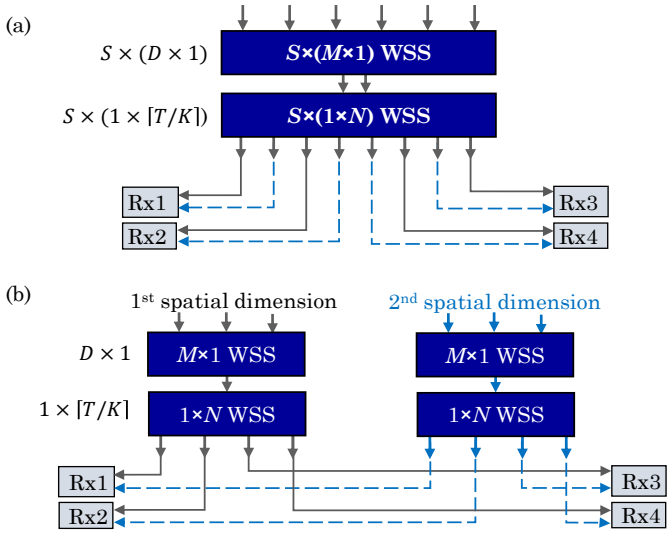


Fig. 10-3. Two A/D module architectures for ROADM colorless directionless operation.

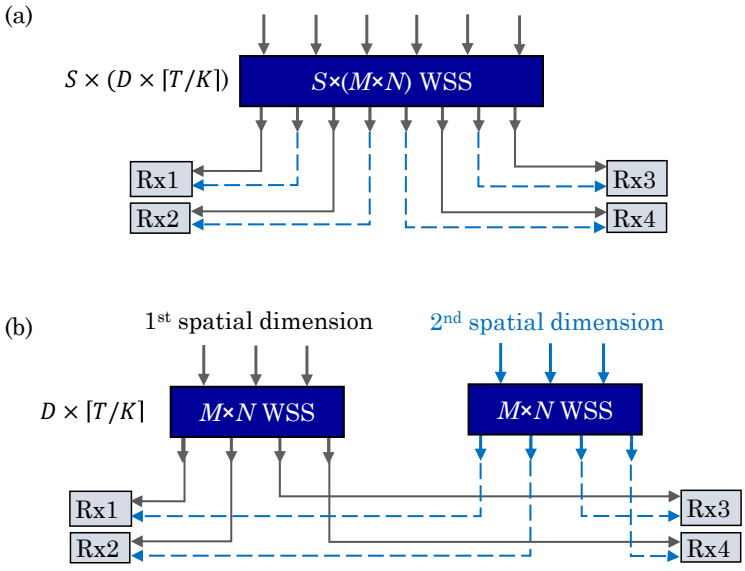


Fig. 10-4. Two M×N WSS-based A/D module architectures for ROADM CDC operation.

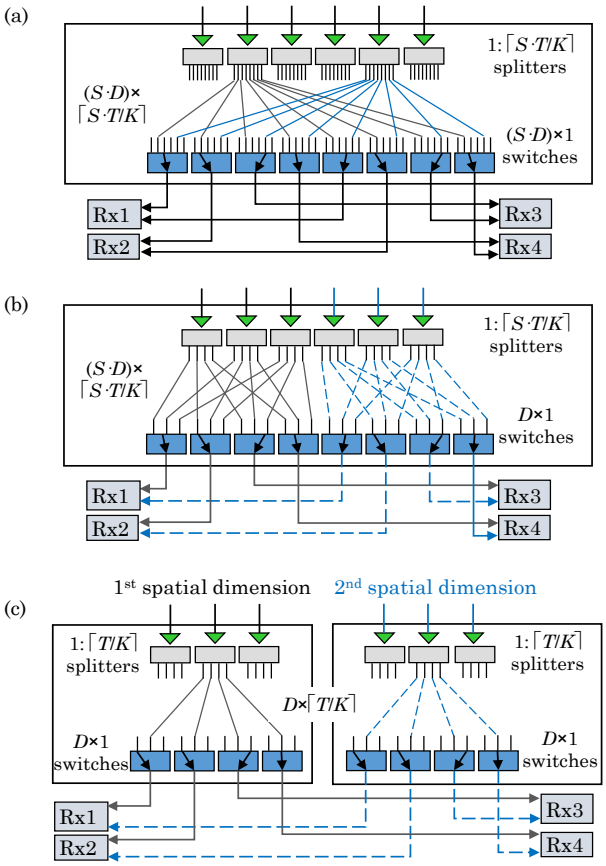


Fig. 10-5. Three MCS-based A/D module architectures for ROADM CDC operation.

Table 10-2 shows the A/D module cost estimation. In the case of CDC operation implemented with MCSs (Fig. 10-5), the cost model is based on the estimation of the number of discrete components (splitters and optomechanical switches), whose costs were obtained by averaging costs from different vendors, and it does not include the cost of packaging and electronics. We consider low-cost single-stage amplifiers with ~17 dB gain (which limits the splitting ratio to about 1:16) and cost C_{EDFA} ~ 0.17. The total cost of amplifiers per A/D module is estimated as $C_{amp}=2 \cdot K \cdot S \cdot D \cdot C_{EDFA}$. The architecture shown in Fig. 10-5(a) supports full CDC switching between all common ports and A/D ports and is required in the case of single-channel TRxs. In Fig. 10-5(b) connectivity is simplified to provide CDC operation per spatial dimension, which is sufficient for integrated Sp-Ch TRxs. In the case of CDC operation with $M \times N$ WSS (Fig. 10-4), we assume WSS configurations with complexity similar to that of a 1×80 WSS, e.g. 8×24. If we decide to use a higher number of common ports M, the number of A/D ports N has to decrease in proportion so that their product remains constant.

Table 10-2 Cost of A/D modules

Corresponding Module	A/D module cost ($C_{A/D}$)	
Fig. 10-3(a)	$2K\left(C_{WSS_{G\times\left(\frac{S}{G}\mathrm{D}\times 1\right)}}+C_{WSS_{G\times\left(1\times\left\lceil\frac{T\cdot S}{G\cdot K}\right\rceil\right)}}\right)$	(10-1)
Fig. 10-3(b)	$2K\frac{S}{G}\left(C_{WSS_{G\times(\mathrm{D}\times 1)}}+C_{WSS_{G\times\left(1\times\left\lceil\frac{T}{K}\right\rceil\right)}}\right)$	(10-2)
Fig. 10-4(a)	$2KC_{WSS_{G\times\left(\frac{S}{G}\mathrm{D}\times\left\lceil\frac{T\cdot S}{G\cdot K}\right\rceil\right)}}$	(10-3)
Fig. 10-4(b)	$2K\frac{S}{G}C_{WSS_{G\times\left(\mathrm{D}\times\left\lceil\frac{T}{K}\right\rceil\right)}}$	(10-4)
Fig. 10-5(a)	$2K\left(S\cdot D\cdot C_{sp_{1:\left\lceil\frac{T\cdot S}{K}\right\rceil}}+\left\lceil\frac{T\cdot S}{G}\right\rceil C_{SW_{(S\cdot D)\times 1}}\right)+C_{amp}$	(10-5)
Fig. 10-5(b)	$2K\left(S\cdot D\cdot C_{sp_{1:\left\lceil\frac{T}{K}\right\rceil}}+\left\lceil\frac{T\cdot S}{G}\right\rceil C_{SW_{D\times 1}}\right)+C_{amp}$	(10-6)
Fig. 10-5(c)	$2K\frac{S}{G}\left(G\cdot D\cdot C_{sp_{1:\left\lceil\frac{T}{K}\right\rceil}}+\left\lceil\frac{T\cdot S}{G}\right\rceil C_{SW_{D\times 1}}\right)+C_{amp}$	(10-7)

10.1.2 Results and Discussions

Based on the cost models described in the previous section, we calculate the total implementation cost of each ROADM architecture. A degree-3 ROADM for an SDM network with six spatial dimensions is considered for the calculations. For the realizations of the designs, we use the components with port count, and relative cost provided in Fig. 10-6 and calculate the Total Relative Cost (TRC) of each design for the three SDM switching paradigms. In order to evaluate the scaling of their implementation cost with the number of A/D ports, we repeat the calculations for 20, 40, and 60 spatial Sp-Ch TRx, each comprising six spatial sub-channels, resulting in 120, 240, and 360 required add and drop ports. To minimize the A/D module cost, we consider a number K of A/D elements in parallel, whose impact on the port count of pass-through WSSs is also taken into account.

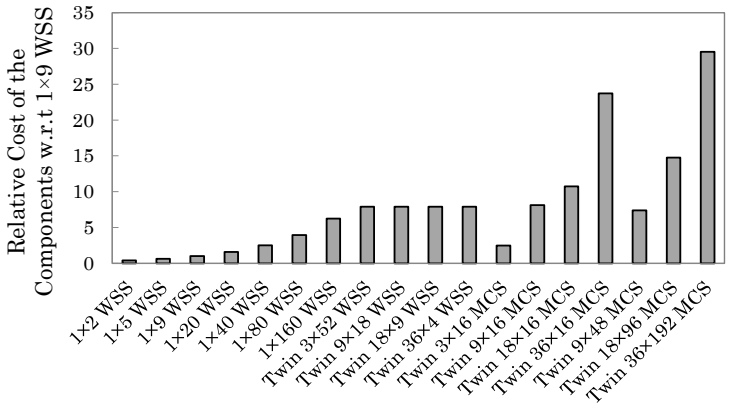


Fig. 10-6. Relative cot of the components w.r.t the cost of 1×9 WSS

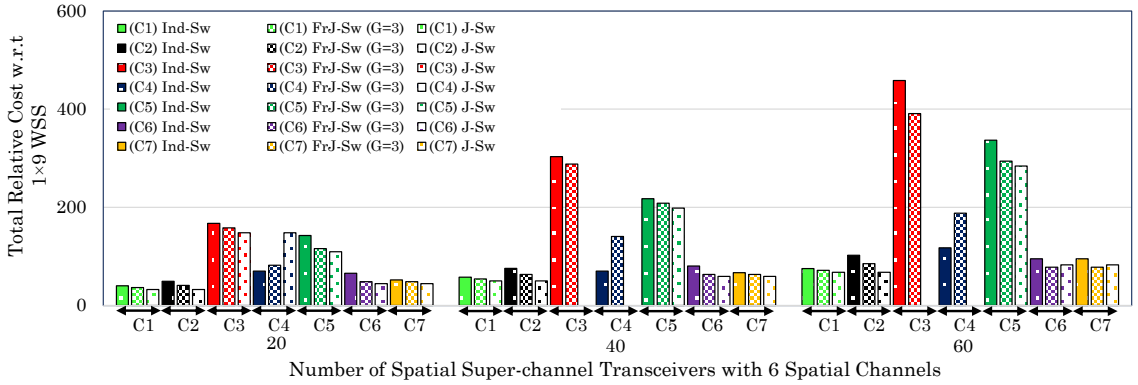


Fig. 10-7. TRC of different ROADM implementations.

Fig. 10-7 shows the resulting TRC values. C1-C7 represent the ROADM architectures with A/D modules illustrated in Fig. 10-3(a-b), Fig. 10-4(a-b), and Fig. 10-5(a-c), respectively.

We found that the most cost-effective architecture is the one, which i) maximizes the number of available A/D ports, and ii) does not heavily increase the port count of pass-through WSSs. Design C7, which utilizes 3×16, 9×48, and 18×96 MCSs (with internal connectivity as shown in Fig. 10-5(c)) for Ind-Sw, FrJ-Sw with G=3, and J-Sw, respectively, is the one that best meets the above conditions and therefore turns out to be the most cost-effective solution. Surprisingly, its implementation cost is very similar to the implementation cost of colorless directionless ROADMs (design C2).

The study also reveals that J-Sw based ROADMs are more cost effective than those implementing Ind-Sw and FrJ-Sw, except for design C4, where the Ind-Sw based realization, by taking advantage of the fact that M×N WSSs with port count 3×52 maximize the available A/D ports with very small increase in the port count of pass-through WSSs, outperforms the other two switching strategies. Design C3,

based on a single $M\times N$ WSS with common ports connected to all spatial dimensions coming from all pass-through WSSs, shows the highest implementation cost, because of the imposed restriction that the $M\times N$ WSS complexity should not be higher than that of a 1×80 WSS. Therefore, the only possible $M\times N$ WSS realization supporting a ROADM of degree 3 in an SDM system with six spatial dimensions is using WSSs with port count 18×9. The very low number of A/D ports supported by this component means that a large number of them is required to support the same number of TRxs that could be supported with fewer components by other architectures (e.g. C4, which can be implemented with $M\times N$ WSS of port count 3×52).

Additionally, we have seen that when the number of parallel A/D elements increases, some of the designs have a huge impact on the port count of pass-through WSSs. These designs turn out to be too costly or impractical to be estimated (e.g. J-Sw based realizations of C3 and C4 when the number of TRxs scales beyond 40) due to the need for WSSs with extremely large port counts.

10.2 The Impact of MIMO Processing on the Performance of SDM Networks

The DSP complexity and the corresponding power consumption may be a showstopper for FMF-based SDM networks. In this section, we investigate the impact on MIMO DSP complexity on the performance and power consumption of FMF-based SDM networks.

It is expected that spatial XT among the modes to be mitigated by MIMO-DSP at the TRx side, with a complexity for real-time implementation determined by the number of spatial channels and their corresponding delay spread [In12] [Ar13], [Ra13]. Current C-form Pluggable (CFP)-based analogue and digital coherent optics modules are focusing on decreasing the cost and volume, as well as the power consumption of the DSP by separating the DSP chip from the module while utilizing smaller Complementary Metal-Oxide-Semiconductor (CMOS) platforms.

While commercial CFP2-ACO DSP modules have been produced utilizing down to 16 nm CMOS platforms [NTT16], it will be extremely difficult to support further power reduction with sub-10 nm CMOS-based DSP modules due to heat dissipation issues [Ki11]. Therefore, even though MIMO-DSP can ideally compensate all the linear impairments in SDM systems, the power consumption of the MIMO-DSP and the overall SDM-TRx can be a limiting factor in the maximum capacity and achievable reach of SDM networks.

In this section, estimation for the power consumption of real-time MIMO-DSP considering the current CMOS technology is presented based on the computation complexity of Frequency-Domain Equalizers (FDE). Also, a reach/power

consumption tradeoff for MIMO-DSP modules is investigated for FMFs with up to 6 spatial modes. Finally, the performance of FMF-based SDM networks with 3 and six modes is compared with SMF-based SDM networks, revealing the impact of power-limited reach on the overall network performance.

10.2.1 Power Consumption of MIMO Processing for SDM

Considering MIMO-FDE with $2S$ spatial and polarization channels, based on training sequence and without any cyclic-prefix, the computational complexity in terms of complex multiplications/bit can be given by [In12], [Ar13],[Ra13]:

$$CC_{FDE} = r_s \frac{2SN_{fft}\log_2(N_{fft}) + 2(2S)^2 N_{fft} + (2S)^3 N_{fft}}{2S(N_{fft} - \left\lceil \frac{T}{T_s} \right\rceil + 1)\log_2(M)} = \tag{10-8}$$

$$r_s \frac{2SN_{fft}\log_2(N_{fft}) + 4SN_{fft} + (2S)^2 N_{fft}}{(N_{fft} - N + 1)\log_2(M)}$$

where, r_s is the oversampling ratio, M the modulation order, T_s the symbol duration, T the channel delay in [sec], $N = \left\lceil \frac{T}{T_s} \right\rceil$ the total channel delay in [symbols], and $N_{fft} = 2^d \geq N$ the Fast-Fourier Transform (FFT) size with d an integer number.

In the above equation, the first term of the numerator accounts for the complexity of 2×FFTs, the second term for the calculation the MIMO channel matrix and the multiplication of its inverse with the received frame, and the third term for the calculation of the inverse of the MIMO channel matrix. Note that, even though the MIMO matrix estimation and inversion is expected to take place every several frames [In12], since here we are focusing on the total power consumption, the worst case was assumed in which all the functions of channel estimation and equalization are performed together. The resulting complexities for 32 GBaud, 100 Gb/s signals over LP FMFs with 3- and 6-spatial modes are compared to the SMF case in Fig. 10-8 for different channel delay spread values. This figure depicts the complexity increment based on the number of spatial channels and the delay spread of the MIMO channel and how for larger delays, larger FFT sizes and number of taps are required for equalization.

In [Sa12], using a reverse-engineering approach, the energy dissipation per real multiplication and per real addition for 90 nm CMOS platforms, was calculated to be 1.5 pJ and 0.5 pJ, respectively. Therefore, considering that a single complex multiplication can be described by four real multiplications and two real additions, an estimation of the power consumption from computational complexity can be drawn.

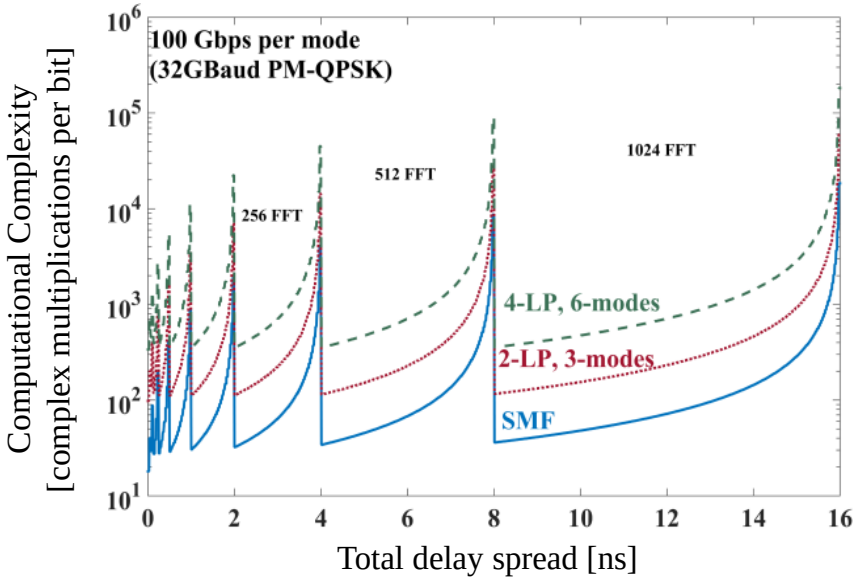


Fig. 10-8. MIMO-DSP complexity

Table 10-3 CMOS power dissipation based on Fig. 2 of [Ki11]

	Moore's law (norm.)	Actual (norm.)	Deviation from Moore's law
90 nm	0.129	0.153	15.7%
45 nm	0.03	0.09	66.7%
22 nm	0.008	0.07	88.6%
15 nm	0.004	0.065	93.9%

In addition, to account for current CFP2-ACO DSP implementations based on sub-20 nm CMOS [Ki11], a power reduction from the 90 nm platforms can be drawn based on dissipation values for different CMOS sizes. Here, we based our analysis on the published results of [Ki11], summarized in Table 10-3. In the first column, the theoretically predicted values based on the Moore's law are shown, while the second column provides values for actual deployed systems. Since the deviation of the actual values from the Moore's law is increasing for smaller sizes, here only the actual values were considered with a deviation of ±5 %.

The resulting power consumptions for 15 nm CMOS, based on the computational complexities of Fig. 10-8, are depicted in Fig. 10-9. Here an FFT size of 1024 was considered. The x-axis, which describes the reach, was converted from the delay spread of Fig. 10-8 considering a chromatic dispersion coefficient of 20 ps/nm/km and a DGD of 0 ps/km, 1 ps/km, and 6 ps/km for the SMF, 2-LP FMF, and 4-LP FMF case, respectively. Particularly, the values for the FMF cases were considered based on recent fabricated fiber links incorporating DGD compensation [Ma14], [Ry13].

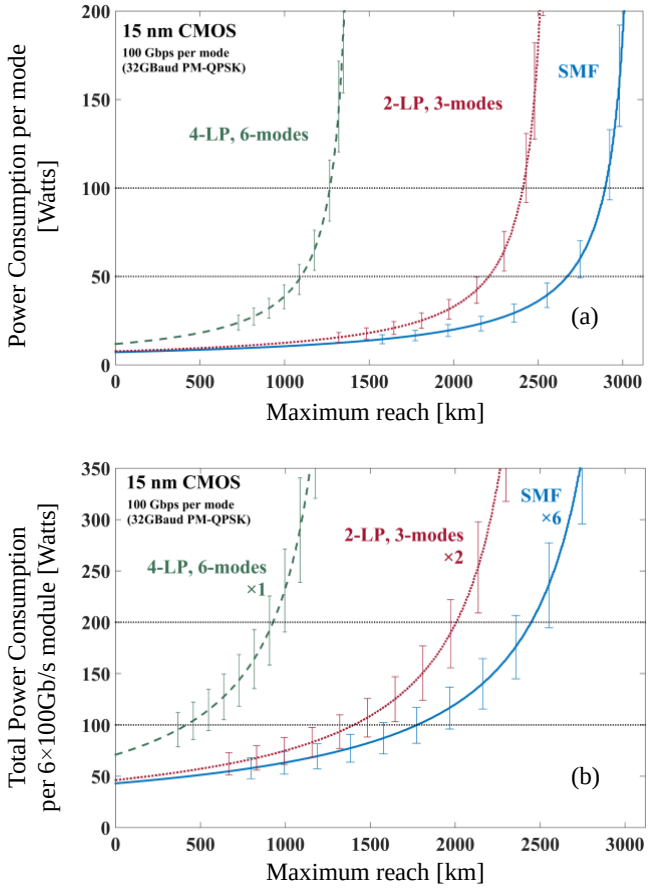


Fig. 10-9. Predicted (a) power consumption per mode and (b) total consumption per 6×100Gb/s module, for 15 nm CMOS.

Particularly, in Fig. 10-9(a) the calculated power consumption per mode is depicted, while in Fig. 10-9(b) the total consumption per 600 Gb/s is depicted based on 6-spatial/parallel channels carrying 100 Gb/s each (i.e. 6×SMFs, 2×3-mode FMFs, and 1×6-mode FMF). Considering that current real-time DSP for long-haul applications consumes 50~100 W per 400~500 Gb/s [Sa16.3] and that current Intel's processors consume powers around 200 W [Intel16], a limit on the maximum reach can be drawn.

The resulting power-limited reaches for different power thresholds and fibers are summarized in Table 10-4 for clarity. In addition, the total power consumption for 750-km, 1000-km, and 1500-km are summarized in Table 10-5. Such estimated distances are of great interest since they are covering most of the deployed European-scale core networks. From this we can conclude that by shifting from parallel SMFs implementations to FMFs with 3- and 6-modes, the realistic power-limited reaches are expected to be reduced by 18~19 % and 60~68 %, respectively, for total consumptions of 100~200 W. Alternatively, the overall consumption increases by 15~18 % and 168~268 %, respectively for the 3- and 6-mode upgrade, between 750-km and 1000-km.

Table 10-4 Power-limited reach based on total consumption per 600 Gb/s module

	100 W	200 W	300 W (50 W/mode)	600 W (100 W/mode)
SMF	2104 km	2609 km	2780 km	2950 km
2-LP FMF	1699 km	2152 km	2306 km	2459 km
4-LP FMF	666 km	1050 km	1180 km	1309 km

Table 10-5 Total power consumption per 600 Gb/s module

	750 km	1000 km	1500 km	fibers × modes
SMF	42.8 W	47.9 W	62.7 W	6 × 1
2-LP FMF	49.1 W	56.7 W	82.2 W	2 × 3
4-LP FMF	112.2 W	176.4 W	-	1 × 6

10.2.2 Network-wide Performance Evaluation

The impact of power consumption on the achievable reach affects the successful establishment of incoming connections. Therefore, to investigate the overall performance reduction due to the power-limited reaches of different MIMO-DSP modules, a networking-wide simulation is performed. The simulation environment is similar to the one described in Section 8.1.1. We use the Spanish national network of Telefónica for our studies as a representation of a typical national-scale backbone transport network encountered in Europe.

For the SDM network evaluations, six spatial dimensions are assumed per link, which can be realized by utilizing either bundles of 6×SMFs, bundles of 2×FMFs with three spatial modes, or FMFs with six spatial modes. Available spectrum equal to 4.8 THz (C-band) on the ITU-T WDM 50 GHz grid per spatial dimension is assumed. Regarding transmission, single-carrier spatial Sp-Ch TRxs with 6 spatial channels, each carrying 32 GBaud DP-QPSK signals have been considered. TRxs with different MIMO-DSP modules for the different fiber types are assumed, resulting in various achievable distances as provided by Table 2. The BP was used as a qualitative performance measure. Note that in our simulator, a blocked connection occurs due to i) transmission distance longer than the optical reach of the signal and ii) unavailability of resources (spectrum, fiber) to establish the connection. The results are summarized in Fig. 10-10. Each data point was obtained by simulating 3×10^5 connection requests and has a confidence interval of 95%. As shown in Fig. 10-10(a), for a consumption of 100 W per TRx module, BuSMFs and 2LP-FMF show similar performance due to the medium-size topology considered, in which most of the connections can be established by the achievable reaches. However, 4LP-FMF shows unacceptable performance resulting in rejecting more than 30% of the incoming connection requests.

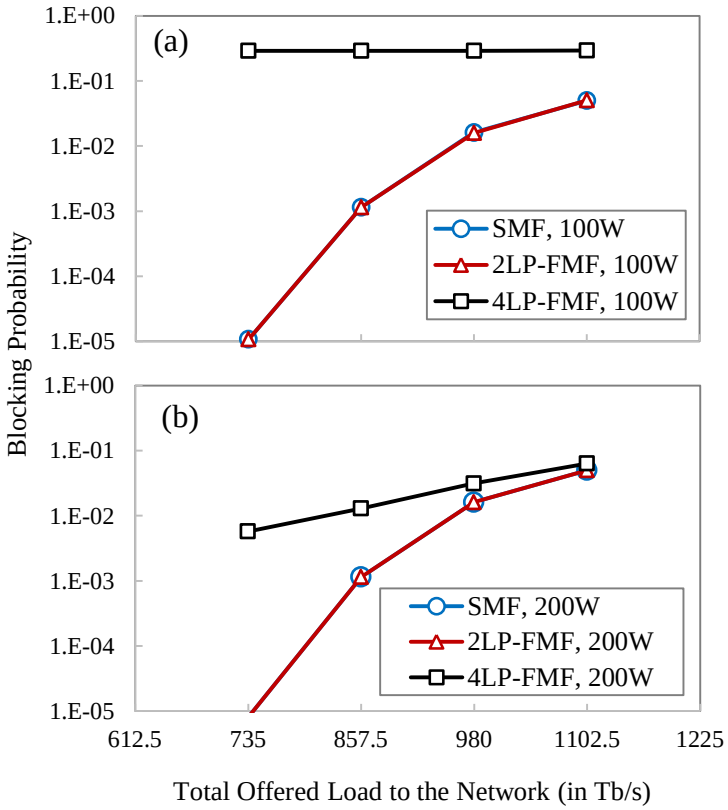


Fig. 10-10. Blocking probability in terms of total offered load considering (a) 100 W and (b) 200 W per 600 Gb/s modules

By increasing the power per module, as shown in Fig. 10-10(b), the performance of 4LP-FMF improves significantly and gets closer to the other cases. The performance for BuSMFs and 2LP-FMFs do not improve much, because the blocking occurs due to the lack of available resource (even when considering 100 W per module). Therefore, it can be concluded that, for a MIMO-DSP module for a 4LP-FMF based SDM network to perform close enough to SDM networks utilizing BuSMFs or 2LP-FMFs, more than two times the power is required.

10.3 **Conclusions**

In this chapter, we investigated the complexity level of spatially integrated components required for the implementation of SDM networks.

First, we analyzed the CD(C) ROADM implementation complexity for the realization of different SDM switching strategies. In this regard, we proposed different colorless directionless and CDC ROADM designs and compared their complexity and implementation cost by calculating their required components and the corresponding configurations. In our studies, we considered various designs based on MCSs, 1×N WSS, and N×M WSSs. We found out that an architecture that

maximizes the number of A/D ports, while keeping the pass-through WSSs port count low, achieves the best cost performance. Moreover, we showed that a CDC design can be as cost-effective as a colorless directionless design based on back-to-back interconnected WSSs.

Second, we analyzed the impact of dense spatial integration on the performance and power consumption of FMF-based SDM networks. We studied the reach/power consumption tradeoff of SDM networks based on the power consumption of MIMO-DSP considering up to 6-spatial channels. We showed that the MIMO-DSP for a 4LP-FMF network requires more than twice the power required for SMF and 2LP-FMF based networks to achieve similar performance.

Chapter 11

Closing Discussion

11.1 Main Contributions

The main contributions of this thesis are:

- The development of several data analytics solutions exploiting the optical spectra of lightpaths to detect and identify soft-failures in filtered optical networks; the proposed solutions have been demonstrated to perfectly detect and identify the most common filter related failures: asymmetric filtering and tight filtering.
- The development of a real-time spectrum surveillance and signal tracking mechanism for filter-less networks; the proposed mechanism has been demonstrated on simulated and experimental data showcasing its outstanding performance to detect and identify Tp failure, as well as to reduce operational margins.
- The development of an autonomic transmission agent, which benefits from SOP traces collected in the coherent receivers to perform local and remote transceiver reconfiguration aiming at enhancing the performance of coherent system for metro networks, while bringing down their cost and power consumption.
- The identification and performance evaluation of different favors of SDM network realization relying on various switching strategies and transmission media. In this regard, the impact of various switching strategies on the performance of SDM networks serving different traffic profiles has been thoroughly studied. Additionally, the benefits and drawbacks of utilizing the three most common transmission media (i.e. bundles of SMFs, MCFs, and FMFs) for the realization of SDM networks have been revealed through a comprehensive set of physical-layer aware performance evaluation. To complement these studies, i) a set of ROADM

complexity analysis for the realization of different SDM switching schemes, and ii) the impact of MIMO-DSP power consumption on the performance of SDM networks utilizing FMFs with different mode-counts are studied.

11.2 List of Publications

11.2.1 Publications in Journals

[JLT16] P. S. Khodashenas, J. M. Rivas-Moscoso, D. Siracusa, F. Pederzolli, **B. Shariati**, D. Klonidis, E. Salvadori, and I. Tomkos, "Comparison of Spectral and Spatial Super-channel Allocation Schemes for SDM Networks," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 34, no. 11, pp. 2710-2716, 2016.

[JLT17] **B. Shariati**, J. M. Rivas-Moscoso, D. M. Marom, S. Ben-Ezra, D. Klonidis, L. Velasco, I. Tomkos, "Impact of Spatial and Spectral Granularity on the Performance of SDM Networks based on Spatial Superchannel Switching," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 35, no. 13, pp. 2559-2568, 2017.

[JLT18] **B. Shariati**, M. Ruiz, J. Comellas, L. Velasco, "Learning from the Optical Spectrum: Failure Detection and Identification [Invited]," IEEE/OSA Journal of Lightwave Technology (JLT), DOI: 10.1109/JLT.2018.2859199, 2018.

[JOCN17] F. Pederzolli, D. Siracusa, **B. Shariati**, J. M. Rivas-Moscoso, E. Salvadori, I. Tomkos, "Improving Performance of Spatially-Switched SDM Optical Networks via Spatial Group Sharing," IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 9, no. 3, pp. B1-B11, 2017.

[JOCN18.1] A. P. Vela, **B. Shariati**, M. Ruiz, F. Cugini, A. Castro, H. Lu, R. Proietti, J. Comellas, P. Castoldi, S. J. B. Yoo, and L. Velasco, "Soft Failure Localization during Commissioning Testing and Lightpath Operation [Invited]", IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 10, no. 1, pp. A27-A36, 2018.

[JOCN18.2] **B. Shariati**, A. Mastropaolo, N. P. Diamantopoulos, J. M. Rivas-Moscoso, D. Klonidis, I. Tomkos, "Physical-Layer Aware Performance Evaluation of SDM Networks based on SMF Bundles, MCFs, and FMFs" IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 10, no 9, pp. 712-722, 2018.

11.2.2 Publications in Conferences

[ACP16] **B. Shariati**, A. Mastropaolo, N. P. Diamantopoulos, J. M. Rivas-Moscoso, F. Pederzolli, D. Siracusa, D. Klonidis, and I. Tomkos, "Spectrally-Spatially Flexible Optical Networking [Invited]," in Proc. of the Asia Communications and Photonics Conference (ACP), 2016.

[ECOC16.1] **B. Shariati**, D. Klonidis, D. Siracusa, F. Pederzolli, J. M. Rivas-Moscoso, L. Velasco, and I. Tomkos, "Impact of Traffic Profile on the Performance of Spatial Superchannel Switching in SDM Networks," in Proc. of the European Conference on Optical Communications (ECOC), 2016.

[ECOC16.2] J. M. Rivas-Moscoso, **B. Shariati**, A. Mastropaolo, D. Klonidis, and I. Tomkos, "Cost Benefit Quantification of SDM Network Implementations Based on Spatially Integrated Network Elements," in Proc. of the European Conference on Optical Communications (ECOC), 2016.

[ECOC18] **B. Shariati**, F. Boiter, M. Ruiz, P. Layec, L. Velasco, "Autonomic Transmission Through Pre-FEC BER Degradation Prediction based on SOP Monitoring," in Proc. of the European Conference on Optical Communications (ECOC), 2018.

[ICTON16] **B. Shariati**, D. Klonidis, J. M. Rivas-Moscoso, and I. Tomkos, "Evaluation of the Impact of Spatial and Spectral Granularities on the Performance of Spatial Superchannel Switching Schemes [Invited]," in Proc. of the International Conference on Transparent Optical Networking (ICTON), 2016.

[ICTON17.1] I. Tomkos, **B. Shariati**, J. M. Rivas-Moscoso, D. M. Marom, and D. Klonidis, "New Frontiers in Optical Communication Networking [Invited]," in Proc. of the International Conference on Transparent Optical Networking (ICTON), 2017.

[ICTON18.1] **B. Shariati**, D. Klonidis, J. Comellas, L. Velasco, I. Tomkos, "Spectrally and Spatially Flexible Optical Networks: Recent Developments and Findings," in Proc. of the International Conference on Transparent Optical Networking (ICTON), 2018.

[ICTON18.2] A. P. Vela, B. Shariati, M. Ruiz, J. Comellas, L. Velasco, "Soft Failure Localization in Elastic Optical Networks," in Proc. of the International Conference on Transparent Optical Networking (ICTON), 2018.

[NOC16] **B. Shariati**, J. M. Rivas-Moscoso, D. Klonidis, D. M. Marom, F. Jiménez, S. Ben-Ezra, P. S. Khodashenas, J. Comellas, L. Velasco, and I. Tomkos, "Options for Cost-effective Capacity Upgrades in Backbone Optical Networks [Invited]," (Invited) in Proc. of the 21st European Conference on Networks and Optical Communications (NOC), 2016.

[OFC16.1] **B. Shariati**, P. S. Khodashenas, J. M. Rivas-Moscoso, S. Ben-Ezra, D. Klonidis, F. Jimenez, L. Velasco, and I. Tomkos, "Evaluation of the Impact of Different SDM Switching Strategies in a Network Planning Scenario", in Proc. of the IEEE/OSA Optical Fiber Communication Conference (OFC), 2016.

[OFC16.2] **B. Shariati**, P. S. Khodashenas, J. M. Rivas-Moscoso, S. Ben-Ezra, D. Klonidis, F. Jimenez, L. Velasco, and I. Tomkos, "Investigation of Mid-term Network Migration Scenarios Comparing Multi-band and Multi-fiber Deployments", in Proc. of the IEEE/OSA Optical Fiber Communication Conference (OFC), 2016.

[OFC17.1] J. M. Rivas-Moscoso, **B. Shariati**, D. M. Marom, D. Klonidis, and I. Tomkos, "Comparison of CD(C) ROADM Architectures for Space Division Multiplexed Networks," in Proc. of the IEEE/OSA Optical Fiber Communication Conference (OFC), 2017.

[OFC17.2] N. P. Diamantopoulos, **B. Shariati**, and I. Tomkos, "On the Power Consumption of MIMO Processing and its Impact on the Performance of SDM Networks," in Proc. of the IEEE/OSA Optical Fiber Communication Conference (OFC), 2017.

[OFC18.1] **B. Shariati,** M. Ruiz, A. Sgambelluri, F. Cugini, L. Velasco, "Real-time Spectrum Surveillance in Filterless Optical Networks," in Proc. of the IEEE/OSA Optical Fiber Communication Conference (OFC), 2018.

[OFC18.2] L. Velasco, **B. Shariati,** A. P. Vela, J. Comellas, and M. Ruiz, "Learning from the Optical Spectrum: Soft-failure Identification and Localization [Invited]," in Proc. of the IEEE/OSA Optical Fiber Communication Conference (OFC), 2018.

[ONDM18] **B. Shariati**, A. P. Vela, M. Ruiz, L. Velasco, "Monitoring and Data Analytics: Analyzing the Optical Spectrum for Soft-Failure Detection and Identification [Invited]," in Proc. of the 24th Conference on Optical Network Design and Modeling (ONDM), 2018.

[OFC19.1] **B. Shariati**, F. Fresi, M. Ruiz, F. Cugini, and L. Velasco, "Optical Signal Tracking for Robust PAM4 Deployment in Filterless Metro Network Scenarios," to be presented at OFC 2019.

[OFC19.2] **B. Shariati**, M. Ruiz, and L. Velasco, "Out-of-field Generic ML Training with In-field Specific Adaptation to Facilitate Deployment," to be presented at OFC 2019.

11.2.3 Book Chapter

[Sh17] **B. Shariati**, J. Comellas, D. Klonidis, L. Velasco, and I. Tomkos, "High Capacity Optical Networks Based on Space Division Multiplexing," in *Provisioning, Recover, and in-Operation Planning in Elastic Optical Networks*, Wiley Series in Optical Communications, ISBN: 978-1-119-33856-7.

11.2.4 Other Works not Included in this PhD Thesis

[OFC19.3] T. Panayiotou, G. Savva, **B. Shariati**, I. Tomkos, and G. Ellinas, "Machine Learning for QoT Estimation of Unseen Optical Networks States," to be presented at OFC 2019.

[APC18.1] D. Marom, M. Blau, J. Macdonald, N. Psaila, C. Sanchez, A. Ellis, J. F. Ferran, F. Jimenez, G. Papastergiou, S. Ben-Ezra, M. Gerola, D. Siracusa, F. Pederzolli, **B. Shariati**, D. Klonidis, I. Tomkos, "Technologies for the Implementation of Spectrally-Spatially Flexible Optical Networks (SS-FON)," presented at APC 2018.

[APC18.2] S. Ben-Ezra, M. Gerola, D. Siracusa, F. Pederzolli, D. Marom, M. Blau, J. Macdonald, N. Psaila, C. Sanchez-Costa, A. D. Ellis, C. Forns, J. F. Jimenez, C. Christodoulia, **B. Shariati**, D. Klonidis, I. Tomkos, "First WDM-SDM Optical Network with Spatial Sub-group Routing ROADM Nodes Supporting Spatial Lane Changes," presented at APC 2018.

[ICC18] G. Savva, G. Ellinas, **B. Shariati**, and I. Tomkos, "Physical Layer Aware Routing, Spectrum, and Core Allocation in Spectrally-Spatially Flexible Optical Networks with Multicore Fibers," presented at the International Conference on Communications (ICC), 2018.

[ICTON17.2] **B. Shariati**, N. P. Diamantopoulos, D. Klonidis, J. Comellas, and I. Tomkos, "On the Benefits of FMF Based Data Center Interconnection Utilizing MIMO-less PAM-M Transceivers [Invited]," presented at the International Conference on Transparent Optical Networking (ICTON), 2017.

[ICTON18.3] G. Savva, K. Manousakis, **B. Shariati**, I. Tomkos, G. Ellinas, "Connection Provisioning in Spectrally-Spatially Flexible Optical Networks with Physical Layer Considerations," presented at ICTON 2018.

[IPS17] **B. Shariati**, D. Klonidis, D. M. Marom, M. Blau, S. Ben-Ezra, M. Gerola, D. Siracusa, J. Macdonald, N. Psaila, C. Sanchez-Costa, A. D. Ellis, J. F. Ferran, F. Jimenez, and I. Tomkos [Invited], "Realizing Spectrally-Spatially Flexible Optical Networks," The Featured Research Highlight Article of IEEE Photonics Society Newsletter, Dec. 2017.

[OFC18.3] L. Gifre, J. L. Izquierdo-Zaragoza, **B. Shariati**, L. Velasco, "Experimental Demonstration of Active and Passive Optical Networks telemetry," presented at the IEEE/OSA Optical Fiber Communication Conference (OFC), 2018.

[OFC18.4] **B. Shariati**, A. Bogris, P. V. Dijk, D. Roeloffzen, I. Tomkos, D. Syvridis, "Free Space Intra-datacenter Interconnection Utilizing 2D Optical Beam Steering," presented at the IEEE/OSA Optical Fiber Communication Conference (OFC), 2018.

[OFC18.5] T. Rokkas, I. Neokosmidis, **B. Shariati**, I. Tomkos, "Techno-economic Evaluations of 400G Optical Interconnect Implementations for Datacenter Networks," presented at the IEEE/OSA Optical Fiber Communication Conference (OFC), 2018.

[ONDM16] F. Pederzolli, D. Siracusa, J. M. Rivas-Moscoso, **B. Shariati**, E. Salvadori, I. Tomkos, "Spatial Group Sharing for SDM Optical Networks with Joint Switching," in Proc. of the 22nd Conference on Optical Network Design and Modeling (ONDM), 2016.

11.3 **List of Research Projects**

11.3.1 European Funded Projects

- FP7 **INSPACE**: Spatial-spectral flexible optical networking: Enabling solutions for a simplified and efficient SDM, 2014-2017.
- H2020 **METRO-HAUL**: Metro high bandwidth, 5G application-aware optical network with edge storage, compute and low latency, 2017-2020.

11.3.2 National Funded Projects

- **SYNERGY**: Service-oriented hYbrid optical NEtwork and cloud infrastructuRe featuring high throuGhput and ultra-low latency. Ref: TEC2014-59995-R, 2015-2017 from MINECO (Ministry of Economy, Industry and Competitiveness)
- **TWINS**: cogniTive 5G application-aware optical metro netWorks Integrating moNitoring, data analyticS and optimization. Ref: TEC2017-90097-R, 2018-2020 from MINECO (Ministry of Economy, Industry and Competitiveness)

11.3.3 Pre-doctoral Funding Scholarship

This thesis has been funded by the European Commission Framework 7 INSPACE research project and by FPI grant supported by Catalan Institution for Research and Advanced Studies (ICREA).

11.4 **Collaborations**

I have spent 18 months, from Oct 2015 - Mar 2017, in Athens Information Technology (AIT), Athens, Greece to carry out research in the framework of INSPACE research project. The topics related to future SDM networks were investigated during this period under the supervision of Dr. Ioannis Tomkos.

In addition, I have collaborated with Nokia Bell Labs, Paris, France in the framework of Metro-Haul project covering the topics explored in goal G.2.

Finally, I have also collaborated with Consorzio Nazionale Interuniversitario per le Telecomunicazioni (CNIT), Pisa, Italy where we have explored some analysis related to goal G.1.

11.5 **Other Achievements**

The outcomes and findings related to the future SDM networks explored in this PhD thesis have been published as the Research Highlight Article of *the IEEE Photonics Society Newsletters, vol. 31, no. 6, issued in December 2017.*

11.6 **Topics for Further Research**

- Some of the optical spectrum monitoring solutions devised in this thesis has been implemented in the CASTOR platform to perform failure localization. CASTOR, standing for Cognitive Architecture to Support Autonomic Network Slicing, is a monitoring and data analytics framework developed in our research group.
- The optical spectrum monitoring solutions proposed for filter failure detection and identification were developed based on simulation measurements in VPI Photonics due to the lack of real measurements. It is very interesting to further extend the work considering experimental or real-field measurements.
- The autonomic transmission agent proposed in the PhD thesis will be experimentally demonstrated to perform local receive configurations

targeting performance improvement of coherent receiver for metro network with lower cost compared to long-haul networks.

- In the context of SDM networks, on the one hand we have investigated the impact of SDM switching constraints on the performance of SDM networks with bundles of SMFs in place to avoid the physical-layer effect of fibers on the performance, and on the other hand, we have performed the physical-aware analysis of SDM networks relying on different transmission media considering just J-Sw scheme in place. It will be valuable to consider both together in a single study; it means to perform a physical-layer aware performance analysis of SDM networks relying on various SDM fibers while every single compatible switching scheme for the fibers under study considered.
- We have considered the same number of spatial dimensions per link and the same type of switching strategies in the nodes while performing the SDM related networking studies; let us call them homogeneous network architectures. However, most probably all the links in a network do not need the same capacity scaling factor over time or all the node do not required the same spatial and spectral granularity levels. Therefore, it is of practical importance to study the performance of heterogeneous network architectures, where the number of spatial dimensions can vary link to link depending on the link load and the spatial granularity of switching schemes can vary node to node depending on the aggregation levels of the traffic. Such studies pave the way for smooth migration from the current networks to the future SDM networks while offer higher cost savings compared to the homogeneous architectures studied in this thesis.
- Even though we have compared the design of different colorless directionless and CDC SDM ROADM architectures in terms of their complexity and implementation cost, a network-wide performance is missing to reveal the benefits and drawbacks of different add/drop configurations in SDM networks.
- Although we have done a MIMO-DSP power consumption analysis and studied its impact on the performance of SDM networks, a comprehensive power consumption analysis of SDM networks is missing from the literature where their corresponding network-wide impact is studied. It is very interesting to consider the impact of power consumption of not only a single configuration of MIMO-DSP, but also other configurations (e.g. individual vs joint MIMO DSP processing) as well as other components like amplifiers on the network-wide performance of SDM networks. The key goal should be defined as finding the best compromise points between the overall network power consumption and the corresponding network-wide performance of different flavors of SDM networks.

List of Acronyms

16QAM	16 Quadrature Amplitude Modulation
8QAM	8 Quadrature Amplitude Modulation
A/D	Add/Drop
ADC	Analogue-to-Digital Converter
ASE	Amplifier Spontaneous Emission
ATA	Autonomic Transmission Agent
BER	Bit Error Rate
BP	Blocking Probability
BPSK	Binary Phase-Shift Keying
BuSMFs	Bundle of Single Mode Fibers
Capex	Capital Expenditure
CASTOR	Cognitive Architecture to Support Telecom clOud Resource sharing
CC	Colorless Contention-less
CD	Chromatic Dispersion
CDC	Colorless Directionless Contention-less
CF	Central Frequency
CFP	C Form-Factor Pluggable
ChBW	Channel Spacing
CMC	Correction Mask Calculator
CMOS	Complementary Metal-Oxide-Semiconductor
CoD	Coherent Detection

DD	Direction Detection
DGD	Differential Group Delay
DP	Dual Polarization
DSP	Digital Signal Processing
DT	Decision Tree
EDFA	Erbium Doped Fiber Amplifier
EON	Elastic Optical Network
ESC	Expected Signal Calculator
FDE	Frequency Domain Equalizer
FEC	Forward Error Correction
FEELING	FailurE causE Localization for optIcal NetworkinG
FeX	Feature Extractor
FFT	Fast Fourier Transform
FMF	Few-Mode Fiber
FM-MCF	Few Mode Multi Core Fiber
FON	Filter-less Optical Network
FPGA	Field Programmable Gate Array
FrJ-Sw	Fractional-Joint Switching
FS	Filter Shift
FSE	Filter Shift Estimator
FT	Filter Tightening
FTE	Filter Tightening Estimator
FWM	Four Wave Mixing
GD	Group Delay
GVD	Group-Velocity Dispersion
HPC	High-Port Count
Ind-Sw	Independent Switching
J-Sw	Joint Switching
LC	Lane Change
LCoS	Liquid Crystal on Silicon
LD	Laser Drift

LP	Linearly Polarized
MAP	Maximum A Posteriori
MCF	Multi Core Fiber
MCS	Multicast Switch
MDA	Monitoring and Data Analytics
MDG	Mode-Dependent Gain
MDL	Mode Dependent Loss
MIMO	Multiple Input Multiple Output
ML	Machine Learning
MMF	Multi-Mode Fiber
MMH	Maximum Margin Hyperplane
MSE	Mean Squared Error
OAA	Observe Analyze Act
OLS	Ordinary Least Square
Opex	Operational Expenditure
OPM	Optical Performance Monitoring
OSA	Optical Spectrum Analyzer
OSNR	Optical Signal to Noise Ratio
PAM4	4-level Pulse Amplitude Modulation
PMD	Polarization Mode Dispersion
PSD	Power Spectral Density
QoT	Quality of Transmission
QPSK	Quadrature Phase-Shift Keying
RMSA	Routing, Modulation, and Spectrum Allocation
ROADM	Reconfigurable Optical Add and Drop Multiplexer
RRC	Root-Raised Cosine
RSA	Routing and Spectrum Allocation
RSCA	Routing, Spectrum, and Core Allocation
RSMSA	Routing, Spectrum, Modulation, and Space Allocation
SDM	Space Division Multiplexing
SDN	Software Defined Network

SG-Sw	Spatial Group Switching
SMF	Single Mode Fiber
SNR	Signal-to-Noise Ratio
SOP	State of Polarization
Sp-Ch	Super-Channel
SPM	Self-Phase Modulation
SSON	Spectrally-Switched Optical Network
SSV	Signal Spectrum Verification
SVM	Support Vector Machine
Tp	Transponder
TRC	Total Relative Cost
TRx	Transceiver
VPI	VPIphotonics: Simulation Software and Design Services
WDM	Wavelength Division Multiplexing
WSS	Wavelength-Selective-Switches
XPM	Cross-Phase Modulation
XT	Crosstalk

List of References

[Ag01] G. P. Agrawal, "Nonlinear Fiber Optics," Academic Press Publisher, 2001.

[An16] C. Antonelli, M. Shtaif, and A. Mecozzi, "Modeling of Nonlinear Propagation in Space-Division Multiplexed Fiber-Optic Transmission," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 36-54, Jan. 2016.

[APV17.1] A. P. Vela, M. Ruiz, and L. Velasco, "Distributing Data Analytics for Efficient Multiple Traffic Anomalies Detection," Elsevier Computer Communications, vol. 107, pp. 1-12, 2017.

[APV17.2] A. P. Vela, M. Ruiz, F. Fresi, N. Sambo, F. Cugini, G. Meloni, L. Potì, L. Velasco, and P. Castoldi, "BER Degradation Detection and Failure Identification in Elastic Optical Networks," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 35, pp. 4595-4604, 2017.

[Ar13] S. Ö. Arik, A. Askarov, J. M. Kahn, "Effect of Mode Coupling on Signal Processing Complexity in Mode-Division Multiplexing," IEEE/OSA Journal of Lightwave Technology, vol. 31, no. 3, pp. 423-431, 2013.

[Ar14] S. O. Arik, J. M. Kahn, and K.-P. Ho, "MIMO Signal Processing for Mode-Division Multiplexing," *IEEE Signal Process. Mag.*, vol. 31, no. 2, pp. 25–34, Mar. 2014.

[Ar15] S. Ö. Arik, K-P Ho, and J. M. Kahn, "Optical Network Scaling: Roles of Spectral and Spatial Aggregation," in Proc. SPIE 9389, Next-Generation Optical Communication: Components, Sub-Systems, and Systems IV, 93890B (2015).

[Ba11] N. Bai, E. Ip, T. Wang, and G. Li, "Multimode Fiber Amplifier with Tunable Modal Gain Using a Reconfigurable Multimode Pump," Opt. Express 19(17), 16601–16611, 2011.

[Bi06] Ch. Bishop, "Pattern Recognition and Machine Learning," Springer-Verlag, 2006.

[Bo17] F. Boitier, V. Lemaire, J. Pesic, L. Chavarría, P. Layec, S. Bigo,E. Dutisseuil, "Proactive Fiber Damage Detection in Real-time Coherent Receiver," in Proc. European Conference on Optical Communications (ECOC), 2017.

[Bo18] F. Boitier and P. Layec, “Automated Optical Networks with Monitoring and Machine Learning,” in Proc. International Conference on Transparent Optical Networks (ICTON), Bucharest, Romania, Jul. 2018.

[Ca12] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, “Modeling of the Impact of Nonlinear Propagation Effects in Uncompensated Optical Coherent Transmission Links," *J. Lightw. Technol.,* vol. 30, no. 10, pp. 1524-1539, May 2012.

[Ca14] J. Carpenter, S. G. Leon-Saval, J. R. Salazar-Gil, J. Bland-Hawthorn, G. Baxter, L. Stewart, S. Frisken, M. A. F. Roelens, B. J. Eggleton, and J. Schröder, “1x11 Few-Mode Fiber Wavelength Selective Switch Using Photonic Lanterns,” Opt. Express, vol. 22, no. 3, pp. 2216-2221, Jan. 2014.

[Ch10] X. Chen and W. Shieh, “Closed-Form Expressions for Nonlinear Transmission Performance of Densely Spaced Coherent Optical OFDM Systems," *Optics Express*, vol. 18, no. 18, pp. 19039-19054, 2010.

[Ch11] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos, "Elastic Bandwidth Allocation in Flexible OFDM-based Optical Networks", J. Lightw. Technol., vol. 29, no. 9, pp. 1354-1366, Mar., 2011.

[Ch16] X. Chen, S. Chandrasekhar, S. Randel, W. Gu, P. Winzer, “Experimental Quantification of Implementation Penalties from Limited ADS Resolution for Nyquist Shaped Higher-Order QAM,” in Proc. Opt. Fiber Comm. Conf. Exhib., Anaheim, CA, USA, 2016, Paper W4A.3.

[Ciena] Whitepaper, “Introducing the Adaptive Network Vision”

[Dh16] M. Dharmaweera, L. Yan, M. Karlsson, E. Agrell, “Nonlinear-Impairments- and Crosstalk-Aware Resource Allocation Schemes for Multicore-Fiber-based Flexgrid Networks,” presented at the European Conf. on Optical Communication (ECOC), 2016, Paper Th.2.P2.SC6.69.

[Do12] C. Dorize, O. Rival, and C. Costantini, “Power Scaling of LDPC Decoder Stage in Long Haul Networks,” in Proc. International Conference on Photonics in Switching (PS), 2012.

[Do15] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu, A. P. T. Lao, “Optical Performance Monitoring: A Review of Current and Future Technologies,” IEEE/OSA Journal of Lightwave Technology (JLT), vol. 34, pp. 525-543, 2015.

[Ei98] T. Eilam-Tzoreff, “The Disjoint Shortest Paths Problem,” in Elsevier Discrete Applied Mathematics (1998), vol. 85, no. 2, pp.113-138.

[El13] A. Ellis, N. Mac Suibhne, F. G. Gunning, and S. Sygletos, “Expressions for the Nonlinear Transmission Performance of Multi-Mode Optical Fiber," *Optics express*, vol. 21, no. 19, pp. 22 834-22 846, 2013.

[Fe12] M. D. Feuer, L. E. Nelson, X. Zhou, S. L. Woodward, R. Isaac, B. Zhu, T. F. Taunay, M. Fishteyn, J. M. Fini, and M. F. Yan, “Joint Digital Signal Processing Receivers for Spatial Superchannels”, Photon. Technol. Lett., vol. 24, no. 21, pp. 1957-1960, 1 Nov. 2012.

[Fe15] M. D. Feuer, “Optical Routing for SDM Networks,” presented at the Eur. Conf. Opt. Commun., Valencia, Spain, 2015, Paper Th.2.5.5.

[Fe16] F. Ferreira, N. Mac Suibhne, C. Sánchez, S. Sygletos, and A. D. Ellis, “Advantages of Strong Mode Coupling for Suppression of Nonlinear Distortion in Few-Mode Fibers," presented at the Optical Fiber Communication Conference (OFC), 2016, Paper Tu2E.3

[Finisar.1] Flexgrid High Resolution Optical Channel Monitor (OCM) [On-line] www.finisar.com, accessed June 2018.

[Finisar.2] Finisar Whitepaper, “Filter Bandwidth Definition of the Waveshaper S-series Programmable Optical Processor,” [On-line] www.finisar.com, accessed June 2018.

[Fo15] N. K. Fontaine, T. Haramaty, R. Ryf, H. Chen, L. Miron, Pascar, M. Blau, B. Frenkel, L. Wang, Y. Messaddeq, S. LaRochelle, R. J. Essiambre, Y. Jung, Q. Kang, J. K. Sahu, S. U. Alam, D. J. Richardson, D. M. Marom, "Heterogeneous Space-Division Multiplexing and Joint Wavelength Switching Demonstration," presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, Mar., 2015, Paper Th5C.5.

[Fo16] N. K. Fontaine, H. Chen, B. Ercan, R. Ryf, G. Labrouille, N. Barré, P. Jian, J. F. Morizur, and D. T. Neilson, “Wavelength Selective Switch with Optimal Steering Element Utilization,” presented at the Opt. Fiber Commun. Conf. Exhib., Anaheim, CA, USA, Mar. 2016, Paper Th5A.6.

[Fu14] S. Fujii, Y. Hirota, H. Tode, and K. Murakami, “On-demand Spectrum and Core Allocation for Reducing Crosstalk in Multicore Fibers in Elastic Optical Networks,” J. Opt. Commun. Netw., vol. 6, no. 12, pp. 1059-1071, 2014.

[Gi18] Ll. Gifre, J.-L. Izquierdo-Zaragoza, M. Ruiz, and L. Velasco, “Autonomic Disaggregated Multilayer Networking,” IEEE/OSA Journal of Optical Communications and Networking (JOCN), vol. 10, pp. 482-492, 2018.

[Gu13] R. Y. Gu, E. Ip, M.-J. Li, Y.-K. Huang, and J. M. Kahn, “Experimental Demonstration of A Spatial Light Modulator-based Few-Mode Fiber Switch for Space-Division Multiplexing,” presented at Frontiers Opt., 2013, Orlando, FL, USA, Paper FW6B.

[Ha09] F. N. Hauske, M. Kuschnerov, B. Spinnler, and B. Lankl, “Optical Performance Monitoring in Digital Coherent Receivers,” J. Lightw. Technol., vol. 26, no. 16, pp. 3623-3631, Aug 2009.

[Ha12] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, “Uncoupled Multi-core Fiber Enhancing Signal-to-Noise Ratio," *Opt. Express*, vol. 20, no. 26, pp. B94-B103, Dec 2012.

[Ho12] K.-P. Ho, “Exact Model for Mode-Dependent Gains and Losses in Multimode Fiber,” *J. Lightw. Technol.,* vol. 30, pp. 3603– 3609, 2012.

[Ho13] K. P. Ho, J. M. Kahn, “Mode Coupling and Its Impact on Spatially Multiplexed Systems,” Optical Fiber Telecommnications VIB, 2013.

[Ho14.1] K. P. Ho, J. M. Kahn, J. P. Wilde, "Wavelength-Selective Switches for Mode-Division Multiplexing: Scaling and Performance Analysis," J. Lightw. Technol., vol. 32, no. 22, Nov., 2014.

[Ho14.2] K. P. Ho and J. M. Kahn, "Linear Propagation Effects in Mode-Division Multiplexing Systems," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 614-628, Feb. 15, 2014.

[Hu15] B. Huang, N. K. Fontaine, R. Ryf, B. Guan, S. G. Leon-Saval, R. Shubochkin, Y. Sun, R. Lingle, and Guifang Li, "All-Fiber Mode-Group-Selective Photonic Lantern Using Graded-Index Multimode Fibers," Opt. Express, vol. 23, no. 1, pp. 224-234, Jan. 2015.

[Hu16] H. Huang, S. Huang, S. Yin, M. Zhang, J. Zhang, W. GuH, "Virtual Network Provisioning over Space Division Multiplexed Optical Networks using Few-Mode Fibers," J. Opt. Commun. Netw., vol. 8, no.10, pp. 726-733, 2016.

[Ik16.1] Y. Ikuma, K. Suzuki, N. Nemoto, E. Hashimoto, O. Moriwaki, and T. Takahashi, "8 × 24 Wavelength Selective Switch for Low-loss Transponder Aggregator," presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, Mar. 2015, Paper Th5A.4.

[IK16.2] Y. Ikuma, K. Suzuki, N. Nemoto, E. Hashimoto, O. Moriwaki and T. Takahashi, "Low-Loss Transponder Aggregator Using Spatial and Planar Optical Circuit," J. Lightw. Technol., vol. 34, no. 1, pp. 67-72, Jan. 2016.

[In12] B. Inan, B. Spinnler, F. Ferreira, D. v. d. Borne, A. Lobato, A. Adhikari, V. A. J. M. Sleiffer, M. Kuschnerov, N. Hanik, S. L. Jansen, "DSP Complexity of Mode-Division Multiplexed Receivers," Optics Express, vol. 20, no. 9, pp. 10859-10869, 2012.

[INSPACE] INSPACE Project website: http://www.ict-inspace.eu/

[Intel16] http://www.intel.com/content/www/us/en/processors/xeon/xeon-phi-detail.html, accessed Oct. 2016.

[Ji16] M. Jinno and Y. Mori, "Unified Architecture and Design Methodology for Integrated SDM-WSS Employing PLC-Based Spatial Beam Transformer Array for Various Types of SDM Fibers," presented at the Opt. Fiber Commun. Conf. Exhib., Anaheim, CA, USA, Mar. 2016, Paper W4B.3.

[Ka15] Kazuhide Nakajima, Pierre Sillard, David Richardson, Ming-Jun Li, René-Jean Essiambre, and Shoichiro Matsuo, "Transmission Media for an SDM-Based Optical Communication System," IEEE Commun. Mag., vol. 53, no. 2, pp. 44-51, Feb. 2015.

[Ki04] D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss, A. E. Willner, "Optical Performance Monitoring," J. Lightw. Technol., vol. 22, no. 1, pp. 294-304, Jan 2004.

[Ki11] D. C. Kilper, G. Atkinson, S. K. Korothy, S. Goyal, P. Vetter, D. Suvakovic, O. Blume, "Power Trends in Communication Networks," IEEE Journal of Selected Topics in Quantum Electronics, vol. 17, no. 2, pp. 275-284, 2011.

[Kl15] D. Klonidis, F. Cugini, O. Gerstel, M. Jinno, V. Lopez, E. Palkopoulou, M. Sekiya, D. Siracusa, G. Thouénon, and C. Betoule, "Spectrally and Spatially Flexible Optical Network Planning and Operations," IEEE Commun. Mag., vol. 53, no. 2, pp. 69-78, Feb., 2015.

[Ko12] M. Koshiba, K. Saitoh, K. Takenaga, and S. Matsuo, "Analytical Expression of Average Power-Coupling Coefficients for Estimating Intercore Crosstalk in Multicore Fibers," IEEE Photonics Journal, vol. 4, no. 5, pp. 1987-1995, Oct 2012.

[Ko16] U. Kozat, G. Liang, K. Kokten, J. Tapolcai, "On Optimal Topology Verification and Failure Localization for Software Defined Networks," IEEE/ACM Transactions on Networking, vol. 24, pp. 2899-2912, 2016.

[Kr15.1] D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-Defined Networking: A Ccomprehensive Survey," in Proc. of IEEE, vol. 103, no. 1, pp. 14-76, Jan 2015.

[Ku17] T. Kupfer, A. Bisplinghof, T. Duthel, C. Fludger, S. Langenbach, "Optimizing Power Consumption of a Coherent DSP for Metro and Data Center Interconnects," in Proc. Opt. Fiber Comm. Conf. Exhib., Los Angeles, CA, USA, Mar 2017, Paper Th3G.2.

[La13] Brett Lantz, "Machine Learning with R," Packt Publishing Ltd, 2nd Edition 2013.

[Li12] Y. Li, L. Gao, G. Shen, L. Peng, "Impact of ROADM Colorless, Directionless, and Contentionless (CDC) Features on Optical Network Performance [Invited]," IEEE/OSA Journal of Optical Communications and Networking, vol. 4, no. 11, pp. B58-B67, 2012.

[Li15] Y. Li, N. Hua, X. Zheng, G. Li, "Capex Advantages of Few-Mode Fiber Networks" presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, Mar. 2015, Paper Th2A.43.

[Lo16] Victor López and Luis Velasco, Elastic Optical Networks: Architectures, Technologies, and Control, Springer, 2016.

[LoVe16] Victor Lopez and Luis Velasco, "Elastic Optical Networks: Architectures, Technologies, and Control," *in the Springer Book Series on Optical Networks*, 2016.

[Ma05] C. Mas, I. Tomkos, and O. Tonguz, "Failure Location Algorithm for Transparent Optical Networks," IEEE Journal on Selected Areas in Communications, vol. 23, pp. 1508-1519. 2005.

[Ma14] R. Maruyama, N. Kuwaki, S. Matsuo, M. Ohashi, "Two Mode Optical Fibers with Low and Flattened Differential Modal Delay Suitable for WDM-MIMO Combined System,", Opt. Exp., vol 22, no. 12, pp. 14311-14321, 2014.

[Ma15.1] D. M. Marom, M. Blau, "Switching Solutions for WDM-SDM Optical Networks," IEEE Comm. Mag., vol. 53, no. 2, pp. 60-68, Feb. 2015.

[Ma15.2] D. M. Marom, J. Dunayevsky, D. Sinefeld, M. Blau, R. Ryf, N. K. Fontaine, M. Monloadiu, S. Randel, C. Liu, B. Ercan, M. Esmaeelpour, S. Chandrasekhar, A. H. Gnauck, S. G. Leon-Saval, J. Bland-Hawthorn, J. R. Salazar-Gil, Y. Sun, L. Grüner-Nielsen, R. Lingle, "Wavelength-Selective Switch with Direct Few Mode Fiber Integration," in Optics Express, vol. 23, no. 5, p. 5723-5737, Feb., 2015.

[Ma17.1] D. M. Marom, P. D. Colbourne, A. D'Errico, N. K. Fontaine, Y. Ikuma, R. Proietti, L. Zong, J. M. Rivas-Moscoso, I. Tomkos, "Survey of Photonic Switching Architectures and Technologies in Support of Spatially and Spectrally Flexible Optical Networking [Invited]," J. Opt. Commun. Netw., vol. 9, no. 1, pp. 1-26, Jan., 2017.

[Ma17.2] Antonia Mastropaolo, "Advanced Techniques for Coherent Transceivers in High Capacity Optical Networks," PhD Thesis, 2017.

[Mi14] P. Mitchell, G. Brown, R. Thomson, N. Psaila and A. Kar, "57 Channel (19×3) Spatial Multiplexer Fabricated Using Direct Laser Inscription," presented at the Opt. Fiber Commun. Conf. Exhib., San Francisco, CA, USA, Mar. 2014, Paper M3K.5.

[Mi16] T. Mizuno, H. Takara, K. Shibahara, A. Sano, and Y. Miyamoto, "Dense Space Division Multiplexed Transmission over Multicore and Multimode Fiber for Long-Haul Transport Systems", J. Lightw. Technol., vol. 34, no. 6, pp. 1484-1493, 15 Mar. 2016.

[Mu13] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear Propagation in Multimode and Multicore Fibers: Generalization of the Manakov Equations," *J. Lightw. Technol.,* vol. 31, no. 3, pp. 398-406, 2013.

[Mu14] A. Muhammad, G. Zervas, D. Simeonidou and R. Forchheimer, "Routing, spectrum and core allocation in flexgrid SDM Networks with Multi-Core Fibers," presented at International Conference on Optical Network Design and Modeling, 2014, Stockholm, Sweden, pp. 192-197.

[Mu15] A. Muhammad, G. Zervas, and R. Forchheimer, "Resource Allocation for Space-Division Multiplexing: Optical White Box versus Optical Black Box Networking", J. Lightw. Technol., vol. 33, no. 23, pp. 4928-4941, 2015

[Na15] K. Nakajima, P. Sillard, D. Richardson, M. Li, R. Essiambre, S. Matsuo, "Transmission Media for an SDM-based Optical Communication System," IEEE Commun. Mag., vol. 53, no. 2, pp. 44-51, Feb. 2015.

[Ne14] L. E. Nelson, M. D. Feuer, K. Abedin, X. Zhou, T. F. Taunay, J. M. Fini, B. Zhu, R. Isaac, R. Harel, G. Cohen, D. M. Marom, "Spatial Superchannel Routing in a Two-Span ROADM System for Space Division Multiplexing," J. Lightw. Technol., vol. 32, no. 4, pp. 783–789, Feb., 2014.

[NTT16] http://www.ntt-electronics.com/en/news/2016/3/industry-first-16nm-100g-200g-coherent-dsp.html, accessed Oct. 2016.

[Pa13] E. Palkopoulou, G. Bosco, A. Carena, D. Klonidis, P. Poggiolini, I. Tomkos, “Nyquist-WDM-based Flexible Optical Networks: Exploring Physical Layer Design Parameters”, J. Lightw. Technol., vol. 31, no. 14, pp. 2332-2339, Jul 2013.

[Pa15] L. Pascar, R. Karubi, B. Frenkel, and D. M. Marom, “Port-Reconfigurable, Wavelength-Selective Switch Array for Colorless / Directionless / Contentionless Optical Add/Drop Multiplexing”, presented at Photonics in Switching 2015, Florence (Italy), pp. 301-303, paper PDP.2.

[Po14] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, F. Forghieri, “The GN-Model of Fiber NonLinear Propagation and its Applications,” J. Lightw. Technol., vol. 32 n.4, pp. 694-721, 2014.

[Po17] Y. Pointurier, “Design of Low-Margin Optical Networks,” *J. Opt. Commun. Netw.*, vol. 9, vo. 1, pp. A9-A17, Jan 2017.

[Pr15] R. Proietti, L. Liu, R. P. Scott, B. Guan, C. Qin, T. Su, F. Giannone, and S. J. B. Yoo, "3D elastic optical networking in the temporal, spectral, and spatial domains," Commun. Mag., vol. 53, no. 2, pp. 79-87, Feb. 2015.

[Pu11] C. Pulikkaseril, L. Stewart, M. A. F. Roelens, G. W. Baxter, S. Poole, S. Frisken, “Spectral Modeling of Channel Band Shapes in Wavelength Selective Switches,” OSA Optics Express, vol. 19, pp. 8458-8470, 2011.

[Ra10] R. Ramaswami, K. N. Sivarajan, G. H. Sasaki, “Optical Networks: A Practical Perspective,” Morgan Kaufmann Publisher, 2010.

[Ra13] S. Randel, P. Winzer, M. Montoliu, R. Ryf, “Complexity Analysis of Adaptive Frequency-Domain Equalization for MIMO-SDM Transmission,” in Proc. 39th European Conference on Optical Communications (ECOC), paper Th.2.C.4, 2013.

[Ra16.1] G. Rademacher and K. Petermann,“Nonlinear Gaussian Noise Model for Multimode Fibers with Space-Division Multiplexing," *J. Lightw. Technol.,* vol. 34, no. 9, pp. 2280-2287, 2016.

[Ra16.2] G. Rademacher, F. Schmidt, K. Petermann, “Optimum Capacity Utilization in Space Division Multiplexed Transmission Systems with Multimode Fibers,” presented at the European Conf. on Optical Communication (ECOC), 2016, Paper Th.2.P2.SC5.62.

[Ra18.1] D. Rafique and L. Velasco, “Machine Learning for Optical Network Automation: Overview, Architecture and Applications,” (Invited Tutorial) IEEE/OSA Journal of Optical Communications and Networking (JOCN), 2018.

[Ra18.2] G. Rademacher, R. Ryf, N. K. Fontaine, H. Chen, R. J. Essiambre, B. J. Puttnam, R. S. Luis, Y. Awaji, N. Wada, S. Gross, N. Riesen, M. Withford, Y. Sun, R. Lingle "Long-Haul Transmission Over Few-Mode Fibers With Space-Division Multiplexing," *J. Lightw. Technol.*, vol. 36, no. 6, pp. 1382-1388, 2018.

[RFC7575] M. Behringer, M. Pritikin, S. Bjarnason, A. Clemm, B. Carpenter, S. Jiang, L. Ciavaglia, "Autonomic Networking: Definitions and Design Goals," IETF RFC 7575, 2015.

[Ri13] D. J. Richardson, J. M. Fini, L. E. Nelson, "Space-Division Multiplexing in Optical Fibres," Nature Photonics, vol. 7, no. 5, pp. 354-362, May, 2013.

[Ro16] C. Rottondi, P. Boffi, P. Martelli, and M. Tornatore, "On the Benefits of Few-Mode Transmission in Ring Metro Optical Networks with Flexible Grid," presented at the Optical Fiber Communication Conference (OFC), 2016, Paper Tu2H.6.

[Ro17] C. Rottondi, P. Boffi, P. Martelli, M. Tornatore, "Routing, Modulation Format, Baud Rate and Spectrum Allocation in Optical Metro Rings with Flexible Grid and Few-Mode Transmission," IEEE/OSA Journal of Lightwave Technology,vol. 35, no. 1, Jan 2017.

[Ry13] R. Ryf, S. Randel, N. K. Fontaine, X. Palou, E. Burrows, S. Corteselli, S. Chandrasekhar, A. H. Gnauck, C. Xie, R-J. Essiambre, P. J. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, L. Gruner-Nielse, R. V. Jensen, R. Lingle J., "708-km Combined WDM-SDM Transmission over Few-Mode Fiber Supporting 12 Spatial and Polarization Modes," in Proc. European Conference on Optical Communication (ECOC), Paper. We.2.D.1, 2013.

[Ry15] R. Ryf, S. Chandrasekhar, S. Randel, D. T. Neilson, N. K. Fontaine, and M. Feuer, "Physical Layer Transmission and Switching Solutions in Support of Spectrally and Spatially Flexible Optical Networks", IEEE Comm. Mag., vol. 53, no. 2, pp. 52-59, Feb. 2015.

[Sa12] S. Savory, "Digital Signal Processing for Coherent Systems," in Proc. Optical Fiber Communication Conference (OFC), paper OTh3C.7., 2012.

[Sa15] G. M. Saridis, D. Alexandropoulos, G. Zervas, and D. Simeonidou, "Survey and Evaluation of Space Division Multiplexing: From Technologies to Optical Networks", Commun. Surveys Tuts., vol. 17, no. 4, pp. 2136-2156, 2015.

[Sa16.1] K. Saitoh, and S. Matsuo, "Multicore Fiber Technology", J. Lightw. Technol., vol. 34, pp.55-66, no. 1, 1 Jan. 2016.

[Sa16.2] Z. Sanjabi Eznaveh, J.E. Antonio Lopez, G. Lopez Galmiche, J. Rodriguez Asomoza, D. Van Ras, P. Sillard, A. Schülzgen, C. M. Okonkwo, and R. Amezcua Correa, "Few Mode Multicore Photonic Lantern Multiplexer, " presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, Mar. 2016, Paper Tu3I.5.

[Sa16.3] S. Savory, "Digital Signal Processing for Multilevel Modulation Formats," in Proc. European Conference on Optical Communications (ECOC), Paper. Tu.2.A.1, 2016.

[Sc12] R. Schmogrow, R. Bouziane, M. Meyer, P. A. Milder, P. C. Schindler, P. Bayvel, R. I. Killey, W. Freude, and J. Leuthold, “Real-Time Digital Nyquist-WDM and OFDM Signal Generation: Spectral Efficiency Versus DSP Complexity”, presented at the 38th Eur. Conf. Exhib. Opt. Commun., 2012, Paper Mo.2.A.4.

[Si14.1] P. Sillard, “Next-Generation Fibers for Space-Division-Multiplexed Transmissions,” presented at the 40th Eur. Conf. Exhib. Opt. Commun., 2014, Cannes, France, Paper Tu.4.1.1.

[Si14.2] P. Sillard, "Scalability of Few-Mode Fibers for Mode-Division-Multiplexed Systems," 2014 IEEE Photonics Conference, San Diego, CA, 2014, pp. 520-521.

[Si17] J. E. Simsarian, P. Winzer, “Shake Before Break per Span Fiber Sensing with In-line Polarization Monitoring,” presented at the Optical Fiber Communications Conference (OFC), 2017.

[Su15] K. Suzuki, Y. Ikuma, E. Hashimoto, K. Yamaguchi, M. Itoh, and T. Takahashi, “Ultra-High Port Count Wavelength Selective Switch Employing Waveguide-Based I/O Frontend”, presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, Mar., 2015, Paper Tu3A.7.

[Su16] K. Suzuki, M. Nakajima, K. Yamaguchi, G. Takashi, Y. Ikuma, K. Shikama, Y. Ishii, M. Itoh, M. Fukutoku, T. Hashimoto, and Y. Miyamoto, “Wavelength Selective Switch for Multi-Core Fiber based Space Division Multiplexed Network with Core-by-Core Switching Capability,” in Photonics Switching (PS), Niigata, Japan, July 2016.

[Th07] R. R. Thomson, H. T. Bookey, N. D. Psaila, A. Fender, S. Campbell, W. N. MacPherson, J. S. Barton, D. T. Reid, and A. K. Kar, "Ultrafast-laser Inscription of a Three Dimensional Fan-out Device for Multicore Fiber Coupling Applications," Opt. Exp., vol. 15, no. 18, pp. 11691-11697, Sep. 2007.

[To17] H. Tode, Y. Hirota, “Routing, Spectrum, and Core and/or Mode Assignment on Space-Division Multiplexing Optical Networks,” J. Opt. Commun. Netw., vol. 9, no. 1, pp. A99-113, Jan., 2017.

[Ve17.1] L. Velasco, A. P. Vela, F. Morales, and M. Ruiz, “Designing, Operating and Re-Optimizing Elastic Optical Networks,” (Invited Tutorial) IEEE/OSA Journal of Lightwave Technology (JLT), vol. 35, pp. 513-526, 2017.

[Ve18.1] L. Velasco, Ll. Gifre, J.-L. Izquierdo-Zaragoza, F. Paolucci, A. P. Vela, A. Sgambelluri*, M. Ruiz, and F. Cugini*, “An Architecture to Support Autonomic Slice Networking [Invited],” IEEE/OSA Journal of Lightwave Technology (JLT), vol. 36, pp. 135-141, 2018.

[Ve18.2] L. Velasco, A. Sgambelluri, R. Casellas, Ll. Gifre, J.-L. Izquierdo-Zaragoza, F. Fresi, F. Paolucci, R. Martínez, and E. Riccardi, “Building Autonomic Optical Whitebox-based Networks,” IEEE/OSA Journal of Lightwave Technology (JLT), 2018.

[VeRu17] Luis Velasco and Marc Ruiz, "Provisioning, Recover, and in-Operation Planning in Elastic Optical Networks", *Wiley Series in Optical Communications*, ISBN: 978-1-119-33856-7.

[Wa11] T. Watanabe, K. Suzuki, T. Goh, K. Hattori, A. Mori, T. Takahashi, T. Sakamoto, K. Morita, S. Sohma, and S. Kamei, "Compact PLC-based Transponder Aggregator for Colorless and Directionless ROADM," presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, Mar. 2011, Paper OTuD3.

[Wi12] P. J. Winzer, "Optical Networking Beyond WDM," IEEE Photonics J., vol. 4, no. 2, p. 647-651, Apr., 2012.

[Wi14] P. J. Winzer, "Making Spatial Multiplexing a Reality," in Nature Photonics, vol. 8, no. 5, pp. 345-348, May, 2014.

[Wi17] P. J. Winzer, D. T. Neilson, "From Scaling Disparities to Integrated Parallelism: A Decathlon for a Decade," IEEE/OSA Journal of Lightwave Technology (JLT), vol. 35, no. 5, March 2017.

[Wo98] T. R. Wolinski, "Polarization in Optical Fibers," in Proc. of the IV International Workshop NOA, 1998.

[Ya12] M. Yan, Z. Tao, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Experimental Comparison of No-Guard-Interval-OFDM and Nyquist-WDM Superchannels," presented at the Opt. Fiber Commun. Conf. Exhib., Los Angeles, CA, USA, 2012, Paper OTh1B.2.

[Ye14] F. Ye, J. Tu, K. Saitoh, and T. Morioka, "Simple Analytical Expression for Crosstalk Estimation in Homogeneous Trench-Assisted Multi-Core Fibers," *Opt. Express*, vol. 22, no. 19, pp. 23 007-23 018, Sep 2014.

[Za13] B. Zaluski, B. Rajtar, H. Habjanix, M. Baranek, N. Slibar, R. Petracic, T. Sukser, "Terastream Implementation of all IP New Architecture," in Proc. 36th Int. Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, May 2013.

[Zh10] 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. Exp., vol. 18, no. 11, pp. 11117-11122, 24 May 2010.