

FEDERAL UNIVERSITY OF PARÁ
INSTITUTE OF TECHNOLOGY
POST GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

**STATISTICAL ANALYSIS AND MARKOV
MODELING OF DYNAMIC RESOURCE
PROVISIONING IN ELASTIC OPTICAL NETWORKS**

Adriana de Nazaré Farias da Rosa

TD - 08/2015

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém-Pará-Brazil
2015

FEDERAL UNIVERSITY OF PARÁ
INSTITUTE OF TECHNOLOGY
POST GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

Adriana de Nazaré Farias da Rosa

**STATISTICAL ANALYSIS AND MARKOV
MODELING OF DYNAMIC RESOURCE
PROVISIONING IN ELASTIC OPTICAL NETWORKS**

TD - 08/2015

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém-Pará-Brazil
2015

FEDERAL UNIVERSITY OF PARÁ
INSTITUTE OF TECHNOLOGY
POST GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

Adriana de Nazaré Farias da Rosa

**STATISTICAL ANALYSIS AND MARKOV
MODELING OF DYNAMIC RESOURCE
PROVISIONING IN ELASTIC OPTICAL NETWORKS**

Doctoral Thesis submitted to examining committee of the Post Graduate Program in Electrical Engineering from Federal University of Pará as a requirement to obtain the title of Doctor in Electrical Engineering.

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém-Pará-Brazil
2015

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Rosa, Adriana de Nazaré Farias da, 1983-

Statistical analysis and markov modeling of dynamic
resource provisioning in elastic optical networks /

Adriana de Nazaré Farias da Rosa. - 2015.

Orientador: João Crisóstomo Weyl

Albuquerque Costa;

Coorientador: Solon Venâncio de Carvalho.

Tese (Doutorado) - Universidade Federal do
Pará, Instituto de Tecnologia, Programa de
Pós-Graduação em Engenharia Elétrica, Belém,
2015.

1. Comunicações óticas - modelos matemáticos.
2. Fibras óticas. 3. Markov, processos de. I.
Título.

CDD 22. ed. 6213827

FEDERAL UNIVERSITY OF PARÁ
INSTITUTE OF TECHNOLOGY
POST GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

STATISTICAL ANALYSIS AND MARKOV MODELING OF DYNAMIC RESOURCE
PROVISIONING IN ELASTIC OPTICAL NETWORKS

AUTHOR: Adriana de Nazaré Farias da Rosa

Doctoral Thesis submitted to examining committee of the Post Graduate Program in Electrical Engineering from Federal University of Pará as a requirement to obtain the title of Doctor in Electrical Engineering.

APPROVED ON 26/06/2015

THESIS EXAMINING COMMITTEE:

Prof. Dr. João Crisóstomo Weyl Albuquerque Costa
(SUPERVISOR – UFPA)

Prof. Dr. Solon Venâncio de Carvalho
(CO-SUPERVISOR – INPE)

Prof. Dr. Carlos Renato Lisboa Francês
(MEMBER – UFPA)

Prof. Dr. Maria Thereza Miranda Rocco Giraldi
(MEMBER – IME)

Prof. Dr. Maria José Pontes
(MEMBER – UFES)

Prof. Dr. Marcelo Eduardo Vieira Segatto
(MEMBER – UFES)

SIGNATURE:

Prof. Dr. Evaldo Gonçalves Pelaes
(Head of the Department – PPGEE/ITEC/UFPA)

UFPA/ITEC/PPGEE

Acknowledgments

It is my pleasure to thank and acknowledge those whose direct involvement made possible the production and completion of this thesis.

I am truly grateful to my supervisors, Professor João Crisóstomo and Professor Solon Carvalho for their time, guidance and support throughout this whole process. I am very much obliged for their most valuable mentorship, advice, encouragement and friendship. All the lessons they have taught me, especially the work attitude, are very precious and for the whole life.

I would like to express my deepest gratitude to Professor Lena Wosinska for accepting me as her Ph.D exchange student and always being very supportive from the start. My technical visit to her group during my Ph.D study has been inspiring me in all these years. Without her excellent supervision, my Ph.D study would not be completed. My very special thanks are conveyed to Dr. Cicek Cavdar, who generously shared her friendship and knowledge with me, and, together with Professor Lena, introduced me to the world of Elastic Optical Networks. In addition, I would like to thank CAPES, under grant agreement n° 0953/11-3, and CNPq within the project n° 870269/2000-3, under grant agreement 142166/2010-3, for the financial support and for providing me the opportunity to study in Sweden.

A special thank you should be given to my “LEA family” for all the happy times we shared during these years. Many thanks for providing me with a joyful environment in my daily work. My infinite gratitude and admiration goes to my closest friends, in especially to my eternal *Vida*, Éder Patrício. They have been an abundant source of happiness and inspiration, and motivation to persevere in the most difficult times.

Living abroad half a world from home is definitely not easy, but I am lucky enough to have found lovely friends in the other side of the ocean. Many thanks go to my “ONLab family”, in especially to my friend Mozhgan Mahloo, for the friendly and joyful relationship we have in our professional and personal lives.

Last but not least, I am forever indebted to my parents and my sister for their everlasting love and indispensable support for my whole life, which has encouraged me to pursue my personal and professional goals. I would like to express my heartfelt appreciation to my beloved Boris Dortschy, for his unconditional understanding and generous encouragement during the last moments of my Ph.D study.

Adriana Rosa

Belém, Pará, Brazil

April, 2015

Abstract

The current trends in optical fiber communications are rapidly approaching the physical capacity limit of standard optical fiber. It is becoming increasingly important to efficiently utilize spectral resources wisely to accommodate the ever-increasing Internet traffic demand. However, the rigid and coarse ITU-T grid specifications regarding the spectrum usage restrict the granularity of bandwidth segmentation and allocation, which frequently causes a mismatch between the allocated and the actual requested link bandwidth. This often leads to over provisioning, where usually more resources are provided than necessary. Recently, the concept of elastic optical networks (EONs) has been proposed in order to reduce the waste of spectra resources. In networks with such feature enabled, modulation parameters and central frequencies are not fixed as in the traditional WDM networks: the resources can be allocated with fine granularity, which can adapt to the granularity of the requested bandwidth without over provisioning. This results in more efficient usage of spectral resources.

However, elastic optical networks must satisfy dynamic connection add and drop over spectral resources that inevitably results in fragmentation of the spectrum. In EONs, spectrum fragmentation is an important and inevitable problem, because it reduces the spectral efficiency. As consequence, the blocking probability (BP) is increased due to scattered gaps in the optical grid. Currently, several metrics have been proposed in order to quantify a level of spectrum fragmentation. Approximation methods might be used for estimating average blocking probability and some fragmentation measures, but are so far unable to accurately evaluate the influence of different requested connection bandwidths and do not allow in-depth investigation of blocking events and their relation to fragmentation.

This thesis presents the analytical study of the effect of fragmentation on requests' blocking probability. In this study, new definitions for blocking that differentiate between the reasons for the blocking events were introduced. An analytical framework based on Markov modeling was proposed in order to calculate steady-state probabilities for the different blocking events and to analyze fragmentation related problems in elastic optical links under dynamic traffic conditions. Statistical investigations were derived in order to investigate how different allocation request sizes contribute to fragmentation and blocking probability.

This work is complemented with the introduction of a new accommodated fragmentation metric that allows better differentiating between very small variations of spectrum occupancy. Moreover, we show to which extent blocking events, due to insufficient amount of available resources, become inevitable and, comparing to the amount of blocking events due to fragmented spectrum, we draw conclusions on the possible gains one can achieve

by system defragmentation. We also show how efficient spectrum allocation policies really are in reducing the part of fragmentation that in particular leads to actual blocking events.

Simulation experiments are carried out showing good match with our analytical results for blocking probability in a small scale scenario. Simulated blocking probabilities for the different blocking events are provided for a larger scale node- and network-wise operation scenario in elastic optical networks.

Keywords: Elastic optical networks. Spectrum allocation. Spectrum fragmentation. Blocking probability. Markov modeling.

List of Figures

Figure 1 – Spectrum utilization for different bit rate links	16
Figure 2 – Time-frequency relation of a windowed sinc-pulse for two different rectangular time window widths	27
Figure 3 – Possible multiplexing scheme of sinc-shaped Nyquist pulses	28
Figure 4 – Time-frequency relation of an OFDM symbol showing three sub-carriers time-domain signals and their spectral relation	30
Figure 5 – Three OFDM channel with three sub-carriers frequency-multiplexed to form a super-channel	31
Figure 6 – Illustration of optical channel assignment for fixed and elastic grid . . .	34
Figure 7 – Network optimization for Elastic Optical Networks	37
Figure 8 – Elastic resource allocation used to carry different services/applications with custom bandwidth allocation	39
Figure 9 – Example of bandwidth selective WXC in the RSA	41
Figure 10 – Example of spectrum fragmentation in a dynamic scenario	43
Figure 11 – Example of spectrum configuration	46
Figure 12 – Possible spectrum configuration resulting from <i>FF</i> , <i>SF</i> , <i>EF</i> and <i>RND</i> SA policies	47
Figure 13 – Flow diagram for the Elastic SA framework	55
Figure 14 – Average fragmentation	65
Figure 15 – Spectrum fragmentation by request size for <i>FF</i> , <i>SF</i> , <i>EF</i> and <i>RND</i> . . .	66
Figure 16 – Total blocking probability for different SA policies	66
Figure 17 – Blocking probability by request size for <i>FF</i> , <i>SF</i> , <i>EF</i> and <i>RND</i>	67
Figure 18 – Fragmentation-blocking for different SA policies and link loads	68
Figure 19 – Resource-blocking for different SA policies and link loads	68
Figure 20 – Contrast between <i>RND</i> , <i>FF</i> , <i>SF</i> and <i>EF</i> SA strategies when resource-blocking starts dominating over fragmentation-blocking	69
Figure 21 – Total blocking probability for different SA policies for large scale EOL .	71
Figure 22 – Fragmentation-blocking for different SA policies and link loads for large scale EOL	71
Figure 23 – Resource-blocking for different SA policies and link loads for large scale EOL	72
Figure 24 – Illustration of a simple <i>broadcast-and-select</i> network topology	73

Figure 25 –Total blocking probability for an elastic network-wise scenario as a function of the load, assuming $N = 300$ slots	74
Figure 26 –Fragmentation-blocking and resource-blocking for an elastic network-wise scenario, assuming $N = 300$ slots; contrast between RND , FF , SF and EF SA strategies when resource-blocking starts dominating fragmentation-blocking	75

List of Tables

Table 1 – Comparison of fragmentation metrics resulting from different SA policies	47
Table 2 – Number of states for <i>Random</i> -SA policy	64
Table 3 – BP of different SA policies for small scale EOL	70
Table 4 – Resource- and fragmentation-blocking of different SA policies for small scale EOL	70
Table 5 – Relation of end-to-end link definition and involved nodes for a network scenario	73

Acronyms

ADC	Analog-to-Digital Conversion
AOTF	Acoustic-Optic Tuneable Filters
AWG	Arrayed Waveguide Gratings
BP	Blocking Probability
BV-OXC	Bandwidth Variable Optical Cross Connect
BV-ROADM	Bandwidth Variable Contentionless Reconfigurable Optical Add/Drop Multiplexer
BV-WSS	Bandwidth Variable Wavelength Selective Switch
CD	Chromatic Dispersion
CDC-ROADM	Colorless, Directionless and Contentionless Reconfigurable Optical Add/Drop Multiplexer
CO-OFDM	Coherent Optical Orthogonal Frequency Division Multiplexing
CTMC	Continuous Time Markov Chain
DAC	Digital-to-Analog Conversion
DSP	Digital Signal Processing
DWDM	Dense Wavelength Division Multiplexing
EDFA	Erbium-Doped Fiber Amplifiers
EOL	Elastic Optical Link
EON	Elastic Optical Network
FFT	Fast Fourier Transform
FGB	Fiber Bragg Gratings
FON	Filterless Optical Network
GBE	Global Balance Equation
IFFT	Inverse Fast Fourier Transform
ILP	Integer Linear Programming
IPTV	Internet Protocol Television
ISI	Inter-Symbol Interference
ITU-T	International Telecommunication Union - Telecommunication Sector
LHS	Left-Hand Side
$LiNbO_3$	Lithium Niobate
LSQR	Sparse Equations and Least Squares

MC-OXC	Multi-Granular Optical-Cross-Connect
MDP	Markov Decision Process
MEMS	Micro-Electro-Mechanical System
MIMO	Multiple-Input-Multiple-Output
MVP	Matrix-Vector Product
N-WDM	Nyquist WDM
OADM	Optical Add/Drop Multiplexer
OAWG	Optical Arbitrary Waveform Generation
OBS	Optical Burst Switching
OOK	On/Off Keying
OPS	Optical Packet Switching
OSNR	Optical Signal-Noise Ratio
OXC	Optical Cross Connect
PLC	Planar Lightwave Circuits
PMD	Polarization Mode Dispersion
PON	Passive Optical Network
PSD	Power Spectral Density
QoT	Quality of Transmission
RAM	Random Access Memory
RHS	Right-Hand Side
ROADM	Reconfigurable Optical Add/Drop Multiplexer
RSA	Routing and Spectrum Allocation
RWA	Routing and Wavelength Allocation
SA	Spectrum Allocation
SDM	Spatial Division Multiplexing
SOA	Semiconductor Optical Amplifiers
SOR	Successive Over-Relaxation
SSS	Spectrum Selective Switch
WBS	Wavelength Band Switching
WDM	Wavelength Division Multiplexing
WRN	Wavelength Routed Network
WSS	Wavelength Selective Switch

Table of Contents

List of Figures	viii
List of Tables	x
Acronyms	xi
1 INTRODUCTION	15
1.1 Motivation of the Thesis	15
1.2 Challenges and Contributions	17
1.3 Outline of the Thesis	20
2 ELASTIC OPTICAL NETWORKING: A CLASSIFIED OVERVIEW	22
2.1 Evolution of Optical Nodes and Networks	22
2.2 State-of-the-Art of Optical Nodes and Networks	24
2.2.1 High Capacity Transmission	24
2.2.2 High Speed Channel Generation	25
2.2.2.1 Nyquist WDM	25
2.2.2.2 Coherent Optical Orthogonal Frequency-Division Multi- plexing (CO-OFDM)	29
2.2.3 Optical Networks that Provide Dynamic Bandwidth	31
2.3 Need for Elastic Spectrum Allocation	32
2.4 Progress toward Elastic Spectrum Allocation	33
2.5 Issues and Challenges in Elastic Optical Networking	34
3 DYNAMIC RESOURCE PROVISIONING IN ELASTIC OPTICAL NET- WORKS	36
3.1 Design Scope Aspects of Networking Optimization	36
3.2 Dynamic Resource Provisioning in EONs: Definition and Complexity	38
3.2.1 Routing and Spectrum Allocation (RSA) Problem	40
3.2.2 Off-line and On-line Approaches	41
3.3 Spectrum Fragmentation Problem	43
3.3.1 Spectrum Allocation (SA) Policies	44
3.3.2 Fragmentation Ratio Calculation	44
4 ELASTIC SA FRAMEWORK BASED ON MARKOV MODELING	49
4.1 Markov Modeling	49
4.1.1 Continuous Time Markov Chains (CTMC)	50

4.1.2	Numerical Methods	51
4.1.2.1	Jacobi Method	53
4.1.2.2	Gauss-Seidel Method	54
4.1.2.3	Successive Over-Relaxation (SOR) Method	54
4.1.2.4	Sparse Equations and Least Squares (LSQR) Method	54
4.2	Elastic SA Framework	55
4.2.1	State-Space Generation	56
4.2.2	Transition-Rate Matrix Generation and Steady-State Probabilities	58
4.3	Calculation of Blocking Probabilities	61
5	STATISTICAL ANALYSIS OF NODE- AND NETWORK-WISE OPERATION SCENARIO IN ELASTIC OPTICAL NETWORKS	63
5.1	Performance Results of Node-Wise Modeling	63
5.1.1	Spectrum Fragmentation Analysis	65
5.1.2	Blocking Probability Analysis	66
5.1.3	Resource- and Fragmentation-Blocking Analysis	67
5.2	Monte Carlo Simulation	69
5.3	Performance Results of Network-Wise Modeling	72
5.3.1	Description of the Network	72
5.3.2	Description of the Simulation	72
5.3.3	Results	74
6	CONCLUSIONS AND FUTURE WORKS	77
6.1	Conclusion	77
6.2	Future Research	78
	Bibliography	81

1 INTRODUCTION

1.1 Motivation of the Thesis

A tremendous growth of Internet traffic volume has been observed in the last few years, mainly due to an increased number of Internet users along with a massive deployment of broadband access networks and the development of new bandwidth demanding Internet applications. With the latest advances in web-based applications available for both business and residential users such as e-services, Internet Protocol television (IPTV), video on demand, cloud and grid computing applications as well as emerging social networks demonstrate unpredictable changes in bandwidth and geographical traffic patterns. Furthermore, low price of mobile services, making it possible to access online applications anywhere and at any time, are also driving the enormous increment and variability of today's demand.

In the near future, Global Internet traffic is expected to continue the increase at a compound annual rate of 34% [1]. As a result of this rapid increase in traffic demands, large-capacity and cost-effective optical fiber transmission systems are required for realizing future optical networks. So far, Wavelength Division Multiplexing (WDM) systems with up to 40 Gb/s capacity per channel have been deployed in backbone networks, while 100 Gb/s interface are now commercially available and 100 Gb/s deployment are expected soon. Moreover, it is foreseen that optical networks will be required to support Tb/s class transmission in the near future. However, scaling to this growth is challenging for conventional optical transmission technology as it suffers from electrical bandwidth bottleneck limitation, and the physical impairments become more severe as the transmission speed increases [2].

On the other hand, the emerging Internet applications call for a more data-rate flexible, agile, reconfigurable, and resource-efficient optical network, while the fixed and coarse granularity of current WDM technology will restrict the optical network to limited bandwidth provisioning, inefficient capacity utilization, and high cost.

The need for cost and energy efficiency as well as scalability requires a flexible network that would have a fine granularity so as to adaptively provide the required capacity to sub- or super-wavelength demands. Approaches such as optical burst switching (OBS) and optical packet switching (OPS) that meet these requirements can only be viewed as long-term solutions since their enabling technologies are not yet mature [3].

To address such issues and efficiently utilize the spectral resources in an optical fiber, elastic optical network (EON) architecture has been recently proposed, as a solution to meet the requirements of future optical networks with today's available enabling technologies [4]. In EON, modulation format and central frequencies are not fixed. To establish

an optical channel each bandwidth variable optical cross connect (BV-OXC) is able to allocate requested spectrum with fine granularity, which allows for efficient spectrum utilization, in contrast to fixed-grid WDM networks.

Recent advances in multi-carrier solutions such as coherent optical orthogonal frequency division multiplexing (CO-OFDM) [5], Nyquist WDM (N-WDM) [6] as well as optical arbitrary waveform generation (OAWG) [7] have set the stage for envisioning fully elastic optical networks. These technologies enable the formation of spectrum-efficient *super-channels*, which consists of several densely packed sub-channel; the optical spectrum is considered to be sliced into a number of frequency slots, with an appropriate width, in order to simplify the network design and modeling [8]. Therefore, offering tunable bit rate from few tens of gigabits per second to terabits per second range. According to the ongoing standardization efforts in ITU-T, the minimum frequency slot unit that can be currently assigned is 12.5 GHz [9]. However, in the future the granularity can be scaled down to 6.25 GHz or below [10].

100 Gb/s-based transmission systems have been commercialized in the recent two years [11]. Since they are compatible with the 50 GHz ITU grid already deployed, the need for replacing the grid did not arise yet. Both the telecom and datacom industries are now considering data rates beyond 100 Gb/s, and 400 Gb/s is receiving a lot of attention as a possible next step. Unfortunately, the spectral width occupied by 400 Gb/s using reasonable modulation formats is too wide to fit in the 50 GHz ITU-T grid, and forcing it to fit by adopting a higher spectral efficiency modulation format would only allow short transmission distances [11]. Figure 1 shows a comparison of an existing ITU-T grid and an elastic optical grid. As one can see, the fixed grid is not able to support bit rates of 400 Gb/s and 1 Tb/s at standard modulation formats, as they overlap with at least one 50 GHz grid boundary [12].

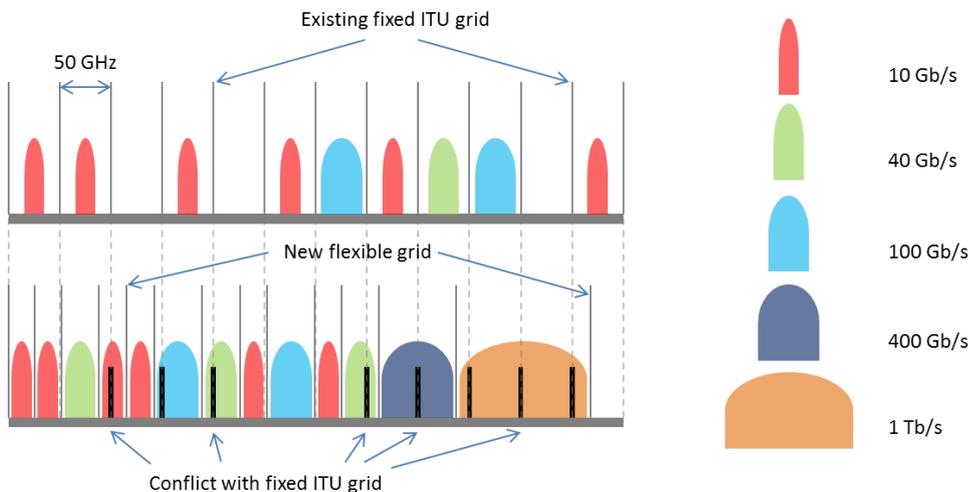


Figure 1: Spectrum utilization for different bit rate links [12].

1.2 Challenges and Contributions

Although elastic optical networks present some similarities to wavelength routed WDM networks, the flexible properties of EONs raise new challenges from the network planning and resource provisioning point of view; in EONs, the allocated spectrum is adapted according to the requested bit rate, transmission distance, and modulation format. Furthermore, the evolution of resource units from an entire wavelength to a frequency slot leads to fundamentally new constraints in spectrum allocation practices, such as e.g., *spectrum contiguity*. Instead of one wavelength, a connection is now assigned one or several frequency slots, depending on whether the capacity requirement of a connection is larger than the capacity of one slot unit. If multiple slots are required, they have to be contiguous, i.e., a connection request can only be satisfied if a sufficient number of free and adjacent slots is available.

The fundamental problem in elastic optical networks is to route and to assign the spectrum resources to accommodate the traffic demands, which is defined as the routing and spectrum allocation (RSA) problem [4]. In the literature, many studies have been conducted during the last three years in order to address the RSA problem [13–17]. The RSA problem in EONs is analogous to the conventional routing and wavelength allocation (RWA) concept in wavelength routed WDM networks. However, due to the new properties of EONs, RWA algorithms cannot be directly applied.

The RSA problem can be divided into two stages for a simpler (suboptimal) solution. As mentioned before, many works have investigated the possible advantages of using a flexible spectrum allocation (SA) concept. Static SA problem is considered for the network design or planning phase [14, 18] where the objective is to minimize the number of necessary frequency slots while provisioning the given traffic demand. Dynamic SA problem, on the other hand, is applied during the network operation, when new lightpaths requests should be served upon arrival [19–21]. Under these conditions, the heterogeneous bandwidth allocation may result in fragmentation of the spectral resources.

Spectrum fragmentation and physical layer impairments, such as polarization and non-linear effects, are further the key issues of EONs. The use of elastic optical channels can result in spectral fragmentation, which increases the blocking probability of connection requests, even the aggregated remaining spectral resources could theoretically provide the capacity, if better utilized, and limits the overall network capacity. In addition, variable bandwidth lightpaths and overlapped sub-carriers in EONs render them more susceptible to physical layer impairments [22–26]. In [22–24], an adaptive quality of transmission (QoT) restoration scheme combining the methods of lightpath re-routing and modulation-format switching has been proposed and experimentally demonstrated to overcome real-time impairments in EONs. In [25], the authors presented an impairment-aware RSA solution for transparent EONs, where signal impairments are compensated by means of

regeneration at selectively placed regenerators. In [26], a QoT-aware RSA algorithm has been proposed, which allocates the modulation level and spectrum by estimating the QoT of the network via a closed-form expression for physical layer impairments.

The statistical analysis of dynamic spectrum allocation was introduced only recently. Blocking probability, resource utilization as well as spectrum fragmentation are the common metrics of performance in EONs, which can be used to dimension network and link capacities while taking into account the dynamics of lightpath requests [8]. It is important to note that the methods utilized to analyze traditional WDM networks are not applicable to assess the performance of EONs. In [27], the authors proposed a birth-death model to analyze the blocking performance for three different spectrum allocation policies, under time-varying traffic in EONs. Although, by applying some relaxation assumptions, the dynamic SA problem is transformed into a static one. In [28], the authors presented an iterative procedure in order to estimate the blocking probability in EONs, with and without spectrum conversion, by approximating the bandwidth utilization ratio. In [8], the authors presented a continuous time Markov chain to calculate spectrum fragmentation and average blocking probability as well as resource utilization of a stand-alone EON link, considering *First-Fit (FF)* and *Random-Fit (RND)* SA approaches. In [29], Beyranvand *et. al.* presented the first attempt for performance evaluation of node- and network-wise operation scenarios in EONs. They developed an analytical framework and investigated average blocking performance and spectrum fragmentation, in scenarios with and without spectrum conversion, considering *FF* and *RND* as spectrum allocation policies. The performance of the framework was compared to exact results for up to eight slots. The authors also show how elastic optical network performance modeling can be broken down to link modeling, in order to significantly reduce the complexity. Recently, in [30] the effect of spectrum fragmentation on the blocking probability of EONs under simplified model of EON operation was investigated with certain fragmentation and utilization. However, due to the high complexity of the model performance, evaluation was only possible by using simulations.

In our study we noticed that the analysis of fragmentation is not sufficient to conclude on the network performance in terms of average BP. The majority of work available in literature tends to assume that blocking probability for future requests should increase monotonically with increasing spectrum fragmentation [20, 31]. Based on this, previous studies have tried to quantify the level of spectrum fragmentation by proposing various fragmentation formulas [31–34]. To the best of our knowledge, none of the previous works on statistical analysis of dynamic spectrum provisioning in the literature considers and investigates the reasons for blocking events, but rather the overall average blocking probability of the system. However, blocking can occur not only because of spectrum fragmentation, but in general due to lack of available resources to serve connection requests. It is usually not considered to which extent and why spectrum allocation policies

proactively minimize the number of blocking events.

In this thesis, we study the spectrum assignment problem in elastic optical networks independently from the routing. Our goal is to provide a mathematical Markov-based model to analyze the statistical performance of elastic optical links for on-line or dynamic traffic conditions. In this respect, the analytical study of the effect of fragmentation on requests blocking probability is presented. The real merits of flexible bandwidth allocation for elastic optical links in terms of wasted (unusable) free spectrum, which is defined as fragmentation, is discussed. Note that fragmentation problem in elastic optical networks is very similar to external memory fragmentation in computer science, so that we borrow some measures and tools from this field in order to quantify fragmentation in elastic optical links.

The major contributions of this work are:

- We formally define the spectrum allocation (SA) problem, taking into account the fragmentation issue. In addition to that, we first present an analytical model to analyze and quantify the blocking probability and fragmentation in an elastic optical link (frequency grid) [35]. Moreover, we compare three existing and one proposed dynamic spectrum allocation schemes by using our framework and discuss the relation of fragmentation and blocking probability [35];
- We also introduce new definitions for blocking that differentiate between the reasons for the blocking events [36]. Furthermore, we introduce an accommodated fragmentation metric from the field of dynamic memory allocation research that allows differentiating between very small variations of spectrum occupancy, which can depend on different spectrum allocation strategies [36]. Based on these new measures, we investigate how various request sizes contribute to fragmentation and blocking probability;
- Moreover, we show to which extend blocking events, due to the fact of insufficient amount of resources, become inevitable and compare to the amount of blocking events due to fragmented spectrum. Based on this comparison, we draw conclusions on the possible gains one can achieve by system defragmentation. We also reveal how efficient dynamic spectrum allocation schemes really are in reducing the part of fragmentation that effectively leads to actual blocking events.
- Finally, we developed a Monte Carlo simulation in order to provide expected results for large scale elastic node- and network-wise operation scenarios for all types of blocking events. Therefore, we show that small scale analytical results can allow drawing conclusions on larger scale elastic grids.

On the other hand, our optical link analysis can give guidance for selection of the right SA policy in the network as well as providing the first step in network-wide evaluation.

Moreover, the model can be easily adopted to evaluate networks based on broadcast-and-select architecture, i.e., without any switching nodes between the source and destination. Examples of such architecture are passive optical network (PON) in the access segment and so called filterless optical network (FON) in the core segment where active switching nodes are replaced by passive splitters/combiners. While different types of PONs are standardized and widely deployed in the fiber access networks, FONs concept has been recently proposed [37] as a cost-effective, energy-efficient and reliable alternative to the active optical switched core networks. The passive *gridless* architecture of FONs makes them naturally suitable for elastic optical networking without the need to replace the switching and filtering devices at nodes. Thanks to the above advantages, we expect that the filterless architecture will gain interest of the network providers in the future optical core and submarine network deployments and we believe that our elastic optical link model will be very useful for analysis of the network-wide performance of elastic FONs.

1.3 Outline of the Thesis

The work included in this thesis addresses important aspects of the resource provisioning paradigm in elastic optical networking, i.e., the performance of dynamic spectrum allocation schemes, spectrum fragmentation, and the effect of fragmentation on requests blocking probability. It is based on research papers published as international contributions in the area of next generation of optical networks. This thesis is organized as follows.

Chapter 2 starts with a summary of the evolution of optical nodes and networks emphasizing the increased flexibility added with each evolutionary stage. It then presents state-of-the-art optical systems scalable to Tb/s, high-capacity transmission experiments and optical systems that provide dynamic bandwidth. The requirements for creating optical networks with such state-of-the-art systems and technology are discussed, leading to the conclusion that there is a need for higher flexibility in the allocation of spectral resources and the flexibility of optical nodes to use functional modules only where required. Ongoing evolution towards elastic spectrum allocation along with several issues brought about by this new paradigm are also presented.

Chapter 3 introduces the concept of elastic allocation of spectral resources, focusing most in dynamic spectrum provisioning and its main matters related to the network level. It then provides an overview of the routing and spectrum assignment (RSA) problem, and discusses one of the key issues in design and optimization of elastic optical networking under dynamic traffic conditions, the so called *spectrum fragmentation* problem. After that, we introduce four spectrum allocation schemes including a genuine one, presented in our conference contribution [35]. With in this contribution, we also introduced for the first time, to the best of our knowledge, an analytical modeling approach for flexible spectrum

allocation under dynamic traffic conditions in the literature. Finally, the fragmentation metrics considered in this thesis are provided, and their ability to reflect the spectrum occupancy is discussed.

Chapter 4, along with chapters 3 and 5, presents the main contributions of this thesis. It first shows the overall process of Markov modeling of real-life systems with inherently stochastic behavior. We then explain the methodology used to evaluate the performance of the model presented in this thesis, including a review of some principal alternatives to efficiently compute the *steady-state* probabilities. Then, the Markov-based framework, presented in our journal contribution [36], is introduced. For the last, we propose new classifications for blocking events that we use in the following to statistically analyze the performance of different SA schemes.

The performance results follow in Chapter 5. After computing the steady-state probabilities of the spectrum usage in an elastic optical link, spectrum fragmentation and blocking probability are statistically evaluated. After that, the analytical results are backed up by results from a Monte Carlo simulation. The results aforementioned have been reported in [36]. Finally, the new definitions of blocking events are applied in the analysis of the dynamic resource provisioning in an elastic optical network scenario.

Finally, Chapter 6 summarizes and evaluates our work, outline ideas for future research activities, and concludes this thesis.

2 ELASTIC OPTICAL NETWORKING: A CLASSIFIED OVERVIEW

The evolution of optical nodes and networks has been characterized by continuous enhancements in capacity, reach, functionality, resiliency, etc. Moreover, such enhancements have inherently meant increasing flexibility. This chapter discusses the evolution of optical nodes and networks and shows how flexibility has been continuously increasing with each evolutionary stage. It also presents state-of-the-art of optical systems that provide high channel and total transmission capacity and discusses their networking requirements. Thus, it shows that there is the need to further increase flexibility in the allocation of spectral resources and the flexibility of optical nodes so that they are able to efficiently support on-demand services and functionality.

2.1 Evolution of Optical Nodes and Networks

Early optical technology was used mainly for point-to-point data transmission, while the processing and networking operation were done in electronics. However, the potential of optical networking to lower costs and increase performance was quickly recognized. One of the first systems proposed for optical networking was based on a passive 8x8 star coupler to broadcast signals from any input port to all output ports [38]. However, the passive star coupler introduced excessive loss, which limited the scalability of the solution.

The advent of WDM introduced additional capacity and flexibility to optical communication systems. Whereas before only a single channel was carried in an optical fiber, now it was possible to transmit data using a number of wavelengths over the same fiber. Two principle pieces of new technology were required for this concept to work on a WDM link: namely optical amplifiers, colored laser sources and multiplexing/demultiplexing devices. For WDM distribution networks, however, additional technology was required. It was necessary to amplify optical signals at intermediate points to compensate for the attenuation introduced by the optical fiber and switching devices. Moreover, a way to add/drop individual wavelengths at intermediate points was needful.

The development of *erbium-doped fiber amplifiers* (EDFAs) in the late 1980s and early 1990s was one of the key enablers of early optical networking. EDFAs also proved cost-effective in point-to-point links as they extended the achievable reach without regeneration [39]. Optical add/drop multiplexers (OADMs) were developed by using a number of technologies, such as acousto-optic tunable filters (AOTF), fiber Bragg gratings (FBG), arrayed waveguide gratings (AWG) and planar lightwave circuits (PLCs). The type of optical networks built using passive OADMs were mainly limited to ring topologies with

OADMs to add/drop wavelengths at each node. Nevertheless, it was now possible to connect one node to various other nodes using the same fiber by selecting different transmission wavelengths for different connections. Hence, ring network topologies were also more flexible than early point-to-point systems.

Although OADMs were widely used for the deployment of ring networks, a different type of optical node was required to implement mesh networks. Thus, in the 1990s several optical cross connect (OXC) [38] designs were developed with the intention of performing the switching function directly in the optical domain, thereby avoiding the use of expensive electronic switching at intermediate nodes. OXCs were first implemented using optical filters and star couplers to construct the switching matrix. Other designs used Lithium-Niobate ($LiNbO_3$) devices or semiconductor optical amplifiers (SOA) as switching gates. In the early 2000s, the micro-electro-mechanical system (MEMS) switches were introduced, providing lower insertion loss and better scalability [39].

However, even OXCs were considered as a feasible solution for small scale systems, the size of optical networks had dramatically grown from the early systems. Optical nodes were required to switch large numbers of wavelengths, e.g., 80 wavelengths in the C-band, from several fibers [39]. Thus, 2D-MEMS switching technology, which scales up to a maximum of around 32 ports, became a bottleneck. This motivated the development of the multi-granular optical cross connect (MC-OXC). MC-OXCs aim to reduce cost and complexity of OXCs by grouping wavelengths together into bands and switching them using single cross connections.

In the hierarchical version of MC-OXC, it is possible to switch the contents of an entire fiber using a single cross-connection (i.e. fiber switching). Alternatively, it is possible to switch bands of channels (i.e. waveband switching) or individual wavelengths (i.e. wavelength switching). Meanwhile, OADMs were further improved by adding reconfigurability in order to select, which wavelength to add/drop at intermediate points. Reconfigurable OADMs (ROADM) [39] were proposed, which made it possible to remotely configure, which wavelengths are added/dropped at an optical node without needing to replace the hardware every time. ROADM designs were further simplified with the introduction of wavelength selective switches (WSS). Here, independent mux/demux and switching devices are not required as WSSs provide both filtering and switching functions. WSS-based ROADMs may have multiple degrees, instead of only two, which makes them usable for the implementation of mesh networks. Further refinements in ROADM design aim to provide colorless, directionless and contentionless operation (CDC-ROADM) [40]. Finally, with the advent of elastic networking, based on the flexibility provided by enabling technologies, e.g., coherent optical OFDM, Nyquist WDM as well as OAWG, an initial bandwidth variable ROADM (BV-ROADM) design has been proposed [4]. BV-ROADMs are based on bandwidth variable wavelength selective switch (BV-WSS) technology, which allows to customize channel bandwidths, thus providing a more flexible and efficient allocation

of spectral resources.

In addition to the obvious advances in capacity and performance, it is clear that with each progressive stage of evolution additional flexibility was introduced, e.g., the introduction of wavelength granularity, the ability to add/drop individual wavelengths, reconfigurability, the increase in number of node degrees, colorless, contentionless and directionless operation and bandwidth variability.

Other types of functionality may also add flexibility to the system, e.g., wavelength, conversion, regeneration, etc. However, in ROADM architectures it is difficult to introduce additional functionality for the through-traffic due to the fact that several wavelengths are simultaneously switched over the same port. On the contrary, OXCs support additional functionality more naturally as wavelengths are split and switched individually. Thus, modules with the required functionality to operate on individual wavelengths can be positioned in the right place within the OXC. However, the requirement for a particular signal processing function is often uncertain, e.g., it may be required for some wavelengths at some time period and for other wavelengths at a different time period. Therefore, modules that provide a per-channel functionality are generally deployed for all wavelengths as there is no possibility of sharing them among several optical paths inside the OXC. A better solution would enable modules to be shared, thus improving modules' utilization and reducing the amount of modules required to satisfy a given demand for the offered functionality, i.e., better hardware utilization and efficiency.

2.2 State-of-the-Art of Optical Nodes and Networks

Current commercial networks are still semi-static as wavelengths are established and left operational for long periods of time. They are also based on the ITU-T grid and implement some kind of ROADM or OXC technology in order to switch wavelengths across the network. In terms of optical network research and experiments, the focus is now on providing high transmission capacity per channel (beyond 100 Gb/s), bandwidth variable transmission, and high total transmission capacity towards Tb/s. This section will review some of the optical channel generation techniques scalable to Tb/s and some of the latest experimental results in this area. Then, it will present bandwidth variable systems and discuss some of the techniques used to achieve bitrate tunability. Next, it will look into the requirements for networking at such high speed and bandwidth variable channels in terms of dynamic resource allocation.

2.2.1 High Capacity Transmission

There are several techniques used or foreseen to extend total capacity, such as increasing the number of channels per fiber, transmitting higher bitrate per channel and using spatial division multiplexing (SDM) in multi-core fibers. In order to accommodate more channels

in a single fiber some experiments make use of the combined spectrum of the C- and L-band [41–47].

Also, multilevel modulation and polarization multiplexing are often used for increasing channel bitrates. For instance, polarization division multiplexing 128-QAM has been reported [42], which yielded a spectral efficiency of 11 bit/s/Hz. Also, SDM has been evolving to support higher numbers of fiber-modes or fiber cores, thereby providing higher transmission capacity. However, SDM also requires complex systems in order to operate and to be compatible with existing infrastructure, e.g., SDM (de-)multiplexers, MIMO processing [48], SDM optical amplifiers [49–51]. In addition to that, a future optical network with SDM transmission, carrying high speed and possibly also legacy channels, would require dynamic, flexible and scalable optical infrastructure.

2.2.2 High Speed Channel Generation

As traffic growth drives the demand for increasingly higher interface rates, research is conducted targeting channels in the terabit per second range. Due to technological limitations, such as the maximum sampling rate of analog-to-digital and digital-to-analog converters (ADC/DAC), approaches that split up the targeted data rate into multiple parallel lower data rate streams have emerged. This leads us to the concept of *super-channels* composed of multiple sub-channels (sub-carriers). There are different approaches that enable the sub-carriers to be efficiently aggregated, such as coherent optical orthogonal frequency division multiplexing (CO-OFDM), Nyquist-WDM (N-WDM) and dynamic optical arbitrary waveform generation (OAWG). In general words, optical OFDM uses orthogonal sub-carriers with spacing equal to multiples of the inverse of the OFDM symbol period [5]. N-WDM uses optical sub-carriers with almost rectangular frequency spectrum closer or equal to the Nyquist limit for *intersymbol-interference-free* transmission. These sub-carriers are multiplexed with spacing closer or equal to the symbol rate with limited inter-sub-carrier crosstalk [6]. The ultimate spectral efficiency is almost identical for both methods under idealized conditions [6]. Finally, OAWG is capable of creating high-bandwidth data waveforms in any modulation format using the parallel synthesis of multiple coherent spectral slices [7].

2.2.2.1 Nyquist WDM

High speed channels are typically build up as *super-channels*. A super-channel can be seen related to or as an evolution of Wavelength Division Multiplexing (WDM), more specifically Dense Wavelength Division Multiplexing (DWDM). A number of N optical carriers are combined coherently to create an aggregated channel that provides N times the capacity of an individual channel.

Above 10 Gb/s, transmission impairments, mainly chromatic dispersion (CD) [38] and polarization mode dispersion (PMD) [38], increase drastically and limit reach and transmission rate. In addition, low-complexity low-cost modulation formats, such as On/Off Keying (OOK), are difficult to realize at high speed, since they require extremely high speed ADCs and digital signal procession (DSP). A way out is to separate the available spectrum in simple channels with a modulation format that allows for small *guardbands*, corresponding to small channels gaps in DWDM.

A required channel gap depends on the decay of sidelobes, out-of-band leakage in real modulation of signals, distortion and windowing effects on transmission impulses. For a maximum packing of simple channels in frequency domain, each channel would (completely) occupy available bandwidth, thus having a rectangular spectrum. Extremely sharp filters would allow to separate the channels and process them individually. Optical filters with a transition bandwidth of approximately only 10% of the pass-bandwidth have been demonstrated in [52].

At the same, the Nyquist-Shannon sampling theorem [53, 54] states that in order not to have inter-symbol interference (ISI) after sampling (A/D conversion), each channel's pulse shape should be such that its periodic extension (by the sampling rate) in frequency-domain does not overlap with the original continuous time signal's spectrum. Both requirements, dense allocation of channels and low ISI after windowing and filtering effects can be achieved with Nyquist-pulses with sinc function characteristics. They possess a (near) rectangular spectrum, enabling data to be encoded in a minimum spectral bandwidth and satisfying by essence the Nyquist criterion of zero ISI. This property makes them very attractive for communication systems, since data transmission rates can be maximized while the spectral bandwidth usage is minimized using only minimal guardbands [55].

The sinc function is defined as:

$$\text{sinc}(x) = \frac{\sin(x)}{x} \quad (2.1)$$

and extends infinitely on the positive and negative time-axis.

For convenience we define

$$\text{rect}(x) = \begin{cases} 1; & |x| < 0.5 \\ 0; & \text{otherwise} \end{cases} \quad (2.2)$$

A perfect rectangular impulse in time domain results in an infinitely wide occupied spectrum due to the relation of symbol length vs. channel spectrum, as

$$\text{rect}\left(\frac{t}{T}\right) \quad \circ\text{---}\bullet \quad T \cdot \text{sinc}(\pi T f), \quad (2.3)$$

one of the technical reasons for having generous channel spacing in fixgrid WDM in order not to distort used high-speed transmission pulses due to filtering effects of, e.g., arrayed

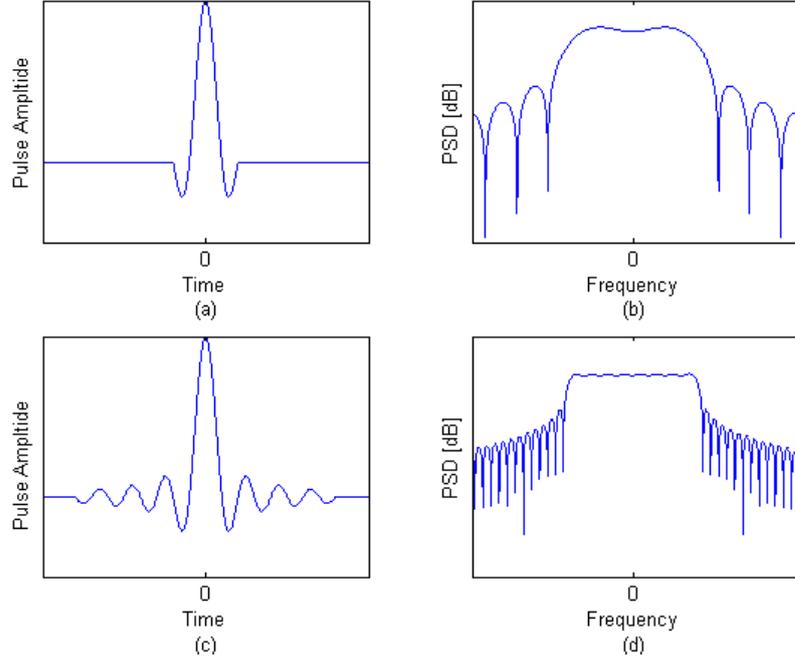


Figure 2: Time-frequency relation of a windowed sinc-pulse for two different rectangular time window widths: (a) $T = 4/F$ with its PSD in (b); and (c) $T = 16/F$ with its PSD in (d) – T and F as in (2.3) and (2.4).

waveguide grating (AWG), if symbol time T gets extremely short, and introduce crosstalk. Fourier theory also tells us, at the same time, the symmetry relation

$$\text{sinc}(\pi Ft) \quad \circ \bullet \quad \frac{1}{F} \cdot \text{rect}\left(\frac{f}{F}\right). \quad (2.4)$$

A signal with a perfectly rectangular spectrum requires generating time-domain impulses ranging from $-\infty$ to $+\infty$ in time, which is not achievable due to the causality requirement in real world. Windowing techniques are applied to each pulse in order to limit its temporal extend and control the roll-off and spectral sharpness in frequency-domain.

Figure 2 illustrates the effect of time-windowing a sinc-pulse (in Fig. 2 (a) and (c)) and the resulting spectrum (in Fig. 2 (b) and (d)). The PSD of a (rectangular) windowed sinc-pulse of temporal expand $k \cdot T$, with T its main-pulse width, can be described as

$$\text{PSD}(f) = \left(kT \text{sinc}(\pi kT \cdot f) \otimes \text{rect}(T \cdot f) \right)^2 \quad (2.5)$$

The operator \otimes denotes the convolution operator. As can be seen from Fig. 2, the wider the window, the more rectangular shaped becomes the pulse spectrum. The two lower figures (Fig. 2 (c) and (d)) assume a time window four times longer ($k = 16$) than the upper pair (Fig. 2 (a) and (b) for $k = 4$). The windowing has no effect on the principal shape of the pulse in time-domain, as long as the pulse-width T is kept constant.

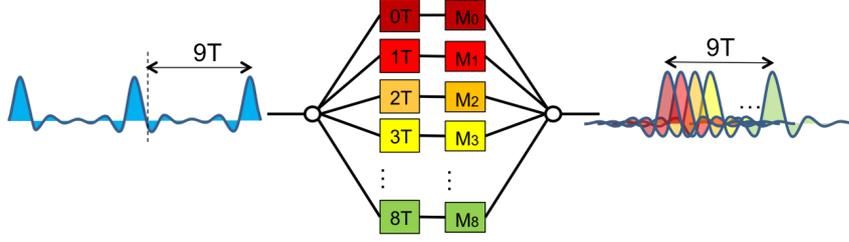


Figure 3: Possible multiplexing scheme of sinc-shaped Nyquist pulses (adapted from [55]).

However, as can be deduced from (2.3), the wider the window, the more confined becomes the left-hand term in the convolution in (2.4). In the limit, i.e. assuming an infinitely wide window ($k \rightarrow \infty$), the left-hand term would resemble a Dirac-pulse, thus not causing any distortion or spectrum widening. The decay and magnitude of sidelobes, as visible in the two spectra in Fig. 2, depends only on the window shape, and can be controlled by using other window types than rectangular. The width of the (near) rectangular spectrum is set by the sinc-pulse width T that also controls the repetition rate in ISI-free N-WDM transmission. Considering the transmitter is flexible and using sufficiently long transmit pulse approximations in time-domain, a transmitter can adjust the pulse transmission rates and utilize available spectrum very efficiently, requiring very reduced guardbands.

A schemes for generating modulated sinc-pulse trains is illustrated in Fig. 3, where we assume, for the sake of simplicity a rectangular window. As in [55], a sinc-pulse train of rate $1/(T \times N)$ is split and fed parallel into N $n \times T$ -delay elements with $n = 0 \dots N$. An information sequence, after serial-to-parallel conversion (not shown), individually modulates each branch's pulse. The output pulses, now each consecutively delayed by T , are combined or time-multiplexed to form a sinc-pulse train of rate $1/T$ and transmitted at high speed. Increasing the number of branches, without increasing incoming pulse rates, allows considering the processing of longer and longer pulses, thus shaping the effective transmit signal spectrum more and more rectangular.

At the receiver side, the high-rate pulse train is split or de-multiplexed by a factor of N to form N pulse trains that can be individually demodulated on lower rate $1/(T \times N)$ again. After parallel-to-serial conversion, the original information sequence can be reconstructed. Fulfilling the Nyquist-ISI criterion, the sinc-shape of each individual pulse guarantees that neighboring pulses do almost not interfere with each other even without wide temporal separation.

Instead of time-delaying a common impulse, using a look-up table large enough to accommodate all possible modulated pulse configuration and their overlap and fast processing in a field-programmable gate array (FPGA) or application-specific integrated circuit (ASIC) can be an all-digital alternative, if ADC can be fast enough.

2.2.2.2 Coherent Optical Orthogonal Frequency-Division Multiplexing (CO-OFDM)

As Nyquist-WDM, coherent optical orthogonal frequency division multiplexing (CO-OFDM) has been recently proposed in response to new challenges in high-speed transmission link for the optical networks. OFDM is a multi-carrier transmission technique where a data stream is simultaneously transmitted over many individually modulated lower-rate sub-carrier tones [2] during the duration of one OFDM symbol. As will be explained, in OFDM, sub-carriers possess an orthogonality property that allows separating the information stream on each sub-carrier after proper demodulation. CO-OFDM combines the advantages of coherent transmission and detection and OFDM modulation and possesses many merits that are critical for future high-speed fiber transmission systems, such as chromatic dispersion and polarization mode dispersion of the transmission system, which can be effectively estimated and mitigated.

In standard multi-carrier modulation, each sub-carrier is demodulated by means of matched filters or correlator principles. Let

$$s(t) = \text{rect}\left(\frac{t - T/2}{T}\right) \cdot \sum_{k=0}^{N_{sc}-1} c_k \exp(j2\pi f_k t) \quad (2.6)$$

be one arbitrary multi-carrier modulation symbol, whereas N_{sc} denotes the number of sub-carriers, T is the OFDM symbol time, c_k would denote the information/modulation symbol and f_k the carrier frequency of the k -th sub-carrier. The rect-function is as defined in (2.2). The matched-filter for demodulation of the k -th sub-carriers (neglecting any channel distortion or delay) would be the conjugate of the sub-carrier pulse

$$m_k(t) = \text{rect}\left(\frac{t - T/2}{T}\right) \cdot \exp(-j2\pi f_k t) \quad (2.7)$$

In OFDM, the sub-carrier spacing, i.e. their relative location in frequency-domain, is in multiples of the inverse of the OFDM symbol time T . That is, the distance between any pair of sub-carrier frequencies fulfills

$$\Delta f_{ij} = f_i - f_j = m \frac{1}{T} \quad (2.8)$$

Applying matched filtering for sub-channel i to such set of sub-carriers results in

$$\begin{aligned} \int_{-\infty}^{\infty} m_i(t) s(t) dt &= \sum_{k=0}^{N_{sc}-1} c_k \cdot \int_0^T \exp(j2\pi \Delta f_{ki} t) dt \\ &= T \sum_{k=0}^{N_{sc}-1} c_k \cdot \exp(j\pi \Delta f_{ki} T) \cdot \text{sinc}(\pi \Delta f_{ki} T) \end{aligned} \quad (2.9)$$

with the sinc-function as defined in (2.1).

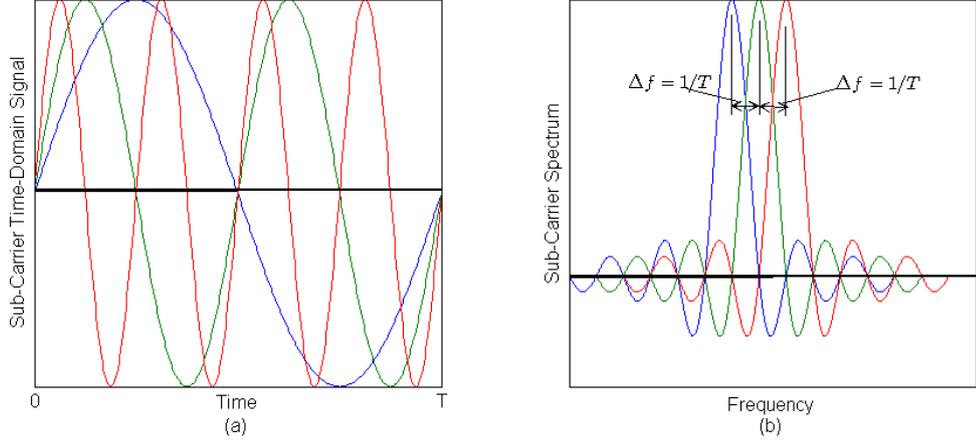


Figure 4: Time-frequency relation of an OFDM symbol showing (a) three sub-carriers time-domain signals and (b) their spectral relation.

Because of (2.8), we have

$$\exp(j\pi\Delta f_{ki}T) \cdot \text{sinc}(\pi\Delta f_{ik}T) = \begin{cases} 1; & i = k \\ 0; & i \neq k \end{cases} \quad (2.10)$$

That is, different sub-carrier are orthogonal to each other. The information symbol c_i of sub-carrier i can be retrieved in (2.9), even without any ISI. This is a major advantage over conventional multi-carrier modulation system with arbitrary sub-carrier spacing, since instead of using filters and requiring significant spacing between sub-carriers, OFDM mitigates inter-carrier interference by proper sub-carrier frequency choice.

Figure 4 shows three sub-carrier signal with frequencies $f_1 = 1/T$, $f_2 = 2/T$, $f_3 = 3/T$. As can be seen, within the OFDM symbol period, all sinusoidal carrier pulses have complete their periods. Their spectra are identical copies of each other, but shifted by multiples of $1/T - 1/T$ and $2/T$ in this case. The spectral shape is determined by the pulse's windowing function. In Fig. 4, the window is assumed rectangular, resulting in sinc-function shape, see also relation (2.3).

The signal processing for sub-carrier generation and modulation in the OFDM transmitter can take advantage of the Inverse Fast Fourier Transform (IFFT). Demodulation in the receiver, equivalent to (2.9) can be implemented as Fast Fourier Transform (FFT). Both FFT and IFFT algorithm have efficient implementations in form of the DFT and IDFT in DSP, if N_{sc} is a power of 2.

Using direct up- and direct coherent down-conversion, the electrical bandwidth requirement can be greatly reduced for the CO-OFDM transceiver, which is extremely attractive for the high-speed circuit design, where electrical signal bandwidth dictates the cost.

For high speed channel generation, OFDM channels can be optically multiplexed to-

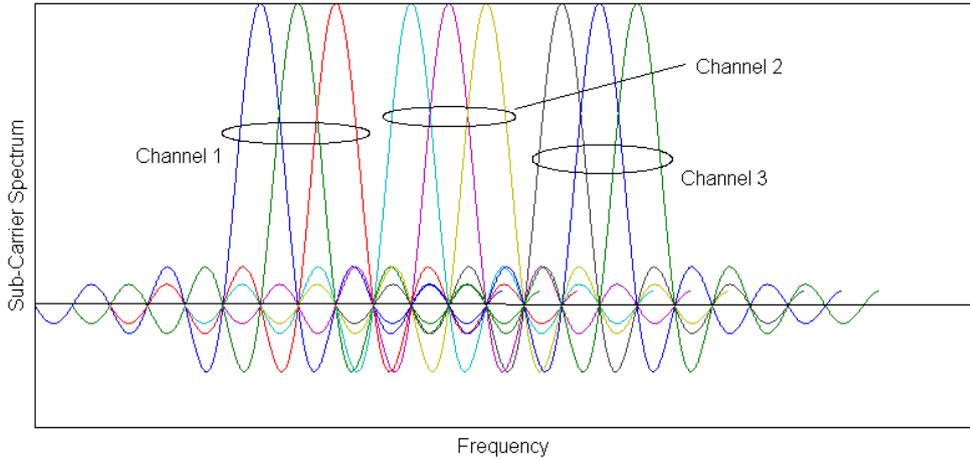


Figure 5: Three OFDM channel with three sub-carriers frequency-multiplexed to form a super-channel.

gether into a *super-channel*, transporting a multiple of the capacity of an individual OFDM channel. This is illustrated in Fig. 5 for a super-channel comprising three OFDM channels with three sub-carriers each. The data stream is first divided into several channels using layer-2 link aggregation, and then OFDM modulated onto optical carriers without or at least minimum *guardbands* in between [2]. In Fig. 5, a guardband of one carrier-spacing is assumed, but would not be necessary in this specific case. The optical carrier frequencies need to be on a grid that preserves orthogonality in between the individual OFDM channels. Carrier drift can be reduced or eliminated by locking all involved optical sources to a common optical comb. This optically aggregated super-wavelength path occupies less spectral resources than the corresponding WDM multiplexing method, thereby leaving room for additional traffic.

2.2.3 Optical Networks that Provide Dynamic Bandwidth

Recently, there have been several demonstrations of dynamic optical bandwidth provisioning whereby transmission bit rates and modulation formats are adapted according to channels' requirements and the impairments of the optical path [56–58]. For instance, channels that traverse longer lightpaths would expect higher degradation, due to impairments, such as CD and PMD, than those traveling only short distances. Therefore, more robust modulation formats may be used for transmission over long lightpaths whereas less robust, more spectrally efficient formats may be used for short distances [56]. This in turn requires flexible provisioning of spectral resources in order to adapt allocated spectrum to channel requirements. The benefits of dynamic bandwidth allocation have been widely studied, from the performance [58, 59] and cost efficiency [57] view points. It provides enhanced resiliency and connectivity as channel transmission distance can be extended in order to reach destinations over alternative paths, although at reduced bit rates. This

may be useful in the case of failure, maintenance, etc. It also makes it possible to improve spectral efficiency, as requested/allocated bitrates can be matched more closely and efficient modulation formats can be used for particular lightpath conditions, e.g. for long reach connections more efficient modulation formats, such as 64-QAM, are employed. In a network with mixed line rates, dynamic bitrate and modulation format transceivers could provide redundancy for several fixed transceivers, e.g. 10 Gb/s, 40 Gb/s, 100 Gb/s. Thus, fewer redundant modules would be required, which enhances cost efficiency [57].

A key enabler of dynamic bandwidth provisioning is the advent of bandwidth variable transceivers, which can adapt their bitrate and modulation format according to requirements. Bitrate tunability can be achieved by adapting the number of optical sub-carriers that make up the high speed multi-carrier channel [56]. Also, single-carrier and multi-carrier systems may adjust modulation formats in order to increase or decrease the number of bits per symbol. Modulation format adaptation by monitoring channel performance and using an automated control plane has been recently demonstrated [22], with alternation between 8-PSK, QPSK and BPSK. Although packing more bits per symbol could provide higher transmission capacity, it also requires higher optical signal-to-noise ratio (OSNR). Therefore, when a particular channel OSNR worsens the control plane triggers a modulation format change to a more robust scheme, which will also usually require more spectrum. On the contrary, when the channel OSNR improves the control plane switches the modulation format back to one that is more efficient and requires less spectrum. Thus, in order to make an efficient use of spectral resources it is necessary to dynamically adapt the allocated bandwidth according to its fluctuating channel requirements.

2.3 Need for Elastic Spectrum Allocation

The dramatic growth of Internet traffic is widely recognized by operators and market analysts alike. This trend is likely to continue due to the emergence of disruptive, bandwidth hungry applications and the proliferation of fiber to the premises and other means of high bandwidth access. In order to keep up with this tremendous growth, operators will need to upgrade their network infrastructure. Currently, the most straightforward and economical way in which this can be done is to deploy additional 10G wavelengths. Thus, new 10G wavelengths are placed 50 GHz or 100 GHz away from other channels, according to the standard ITU WDM grid, until the available bandwidth is exhausted. However, the maximum capacity that can thus be provided is 800 Gb/s (i.e. 80x10G using only the C-band), which is already insufficient for heavily used backbone network links [11]. Furthermore, providing additional capacity in this manner is highly inefficient in terms of the spectral resources that are consumed.

The immediate solution to this problem is to deploy 100G links, despite their higher cost compared to 10x10G. 100G is more spectrally efficient than 10G as it can still fit

in a 50 GHz channel bandwidth, thanks to the use of DP-QPSK modulation, coherent detection, extensive use of forward error correction and electronic impairment mitigation. However, this solution is expected to be viable only for a limited time, while the Internet traffic growth is leading to a strong requirement for 400 Gb/s and 1 Tb/s in the long term. In addition to that, bitrates beyond 100G are unlikely to fit in a 50 GHz channel as that would require more complex multilevel modulation formats with higher OSNR requirements and consequently shorter reach.

Experimental demonstrations of optical transmission systems scalable to Tb/s have required more bandwidth than that provided by a single DWDM channel and, in some cases, not even multiples of 50 GHz. These channels are not supported by current optical network infrastructure designed to conform to the standard 50 GHz or 100 GHz grid. For example, OXCs and ROADMs allocate only discrete 50 GHz slots of bandwidth due to their internal WDM (de-)multiplexers. Channels that require wider bandwidths are severely distorted, if passed through such devices. Therefore, in order to efficiently support high speed channels flexible bandwidth infrastructure is required.

Elastic spectrum allocation can also increase efficiency when transporting legacy channels. For example, 10G channels can be deployed with a 25 GHz channel spacing, thereby doubling spectral efficiency [60]. Narrower channel spacings may be used for lower channel bit rates, e.g., 12.5 GHz for 2.5 Gb/s [61]. Another advantage is the support for bandwidth variable transceivers, which can provide dynamic bandwidth, achieve cost reductions as fewer single type transceivers may be required, and trade-off reach and spectrum usage. For example, consider a transmitter that needs to switch from 8-PSK to QPSK in order to increase reach or maintain connectivity. In order to maintain the original bitrate, the new QPSK channel will require additional bandwidth as it is less spectrally efficient than 8 PSK. In this way, elastic spectrum allocation can be used to adjust allocated bandwidths according to channels' requirements.

2.4 Progress toward Elastic Spectrum Allocation

The first demonstration of an elastic optical network, based on OFDM transmission, was presented in [62]. Since then, a number of studies have investigated elastic networking showing significant gains in network mean traffic [63], required spectral resources [56], capacity [59] and cost [57]. Also, several demonstrations of multi-bitrate transmission over long distances show that it is feasible for multi-Tb/s and lower speed channels to coexist in the same link using flexible spectrum allocation [64]. In the last two years, there have been important results showing the feasibility of automated adaptive transmission [65] and networking [66]. Other works have developed aspects such as translucent elastic regeneration [62] and dynamic failure restoration [67] for elastic optical networks.

Recently, ITU-T has extended its recommendation G694.1 [9], to include the concept

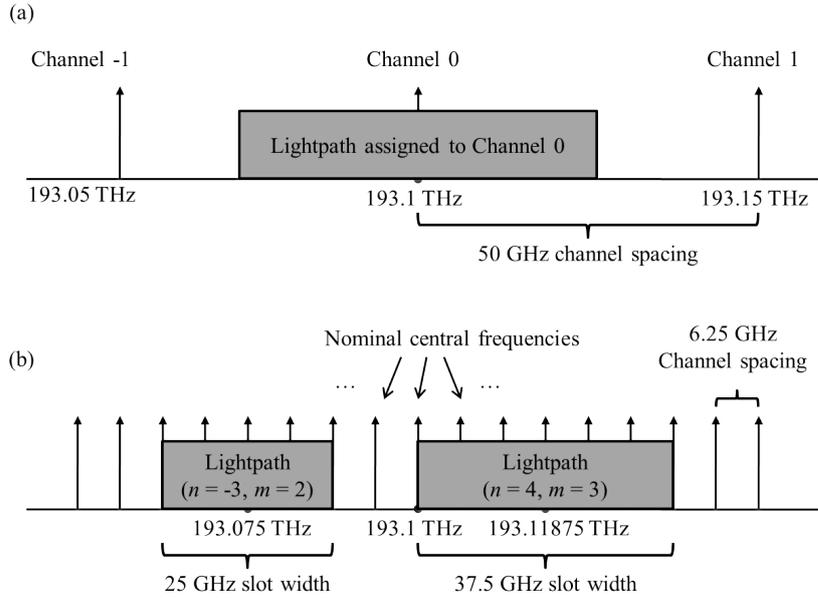


Figure 6: Illustration of optical channel assignment for (a) fixed and (b) elastic grid (adapted from [13]).

of elastic grid. A new DWDM grid has been developed within the ITU-T Study Group 15 by defining a set of nominal central frequencies, channel spacing, and the concept of frequency slot. The main difference with respect to the WDM network is the way the basic unit of switching is identified which is the frequency slot now rather than a wavelength. A frequency slot is defined by its nominal central frequency in the whole spectrum range and its slot width. The set of nominal central frequencies can be built using the following expression $f = 193.1 + n \times 0.00625$ THz, where 193.1 THz is ITU-T “anchor frequency” for transmission over the C-band, and n is a positive or negative integer including 0. It means that the central frequency can be moved in the C-band at 6.25 GHz steps. The slot width determines the “amount” of optical spectrum regardless of its actual position in the spectrum. A slot width is constrained to be $m \times 12.5$ GHz, where m is an integer greater than or equal to 1 and 12.5 GHz, since an even number of 6.25 GHz slots has to be allocated around the central frequency, as illustrated in Fig. 6.

Although the progress toward elastic spectrum allocation are remarkable, in order to develop truly elastic infrastructure much work on areas such as elastic transceivers, optical node design, routing and spectrum allocation algorithms, network control as well as management is still required.

2.5 Issues and Challenges in Elastic Optical Networking

Data plane architectures able to allocate spectrum flexibly are fundamental for elastic optical networks. Also, if flexible allocation is required for other kinds of network resources, e.g., time and space, such functionality needs to be supported by the data plane. In addition, a desirable feature of optical nodes is to facilitate smooth hardware migration

from fixed to flex-grid. Thus, it would be possible to carry legacy signals with low cost in elastic hardware while using the more expensive elastic hardware only where required.

Also, in the area of flexible transponders further developments are required. There have been several demonstrations of bitrate-variable transmitters where the number of sub-carriers or the modulation format are adapted to achieve the desired bitrate and spectral efficiency [68, 69]. However, such transceivers can generate only a single channel with a specific bitrate. Therefore, when high-speed flexible transponders operate in low speed modes, part of their transmission capacity is wasted. This has led to the idea of sliceable transceivers, which are able to slice their total capacity for transmission towards different destinations. Transceiver capacity may be sliced in one or several dimensions, e.g., frequency, time, space and polarization. However, up until now only the frequency domain has been exploited to slice a transceiver's capacity.

Elastic optical networking also presents new challenges to the control and management planes. For instance, WDM networks utilize routing and wavelength allocation (RWA) algorithms to find available resources for new requests. If wavelength conversion is not permitted, the same wavelength needs to be available from source to destination, i.e. wavelength continuity constraint. In EONs, similar algorithms are required to allocate routes and spectrum. Here, however, the problem is more difficult due to the new flexible spectrum allocation, where elastic spectrum bands rather than single wavelengths are considered. For new requests with specific bandwidth requirements, routing and spectrum allocation (RSA) algorithms need to identify sufficiently wide spectrum slots that are available from source to destination. Furthermore, as channels are added and removed, they leave behind non-contiguous slots of free spectrum. Although these fragments may add up to a considerable amount of bandwidth, new channel requests may be blocked due to the lack of sufficient contiguous spectrum [70]. Spectrum fragmentation may be prevented to some extent by introducing appropriate policies in RSA algorithms. Alternatively, techniques to defragment the spectrum may also be utilized to periodically re-optimize the network resources. In the next chapter, we will discuss in more details issues involving the design and optimization of elastic optical networks.

3 DYNAMIC RESOURCE PROVISIONING IN ELASTIC OPTICAL NETWORKS

Advances in physical layer technologies, as shown in Chapter 2, couple with enhanced control plane solutions provide the basis to optimize the overall performance on the network level. Aspects as strategies of survivability for elastic optical networking, to enable high resiliency against network failures, new network control and management schemes, network virtualization as well as energy efficiency strategies, which are vital to facilitate operation and maintenance of elastic optical network, need to be studied. Furthermore, efficient resource management strategies are required for the network planning and especially for dynamic spectrum allocation. In the following chapter, we discuss issues involved in the optimization of elastic optical networks, focusing most of dynamic spectrum provisioning, spectrum fragmentation and some elastic spectrum assignment schemes, which are common used to efficient allocate the spectral resources in EONs.

3.1 Design Scope Aspects of Networking Optimization

The emergence of elastic optical networking as a method to increase resource efficiency and provide advanced functionalities poses significant challenges on the networking level. The introduction of extra degrees of flexibility dictates the enhancement of the network planning and optimization procedure to additionally consider the manner in which these extra parameters should be set. Aspects such as network performance (examined on the connection level and on the network level), cost (consisting of capital and operational expenditures), energy efficiency, and control plane requirements are usually considered as research areas of interest, and have been attracted not only universities, but also important companies in the area of Telecommunication, e.g., Alcatel-Lucent, Fujitsu Network Communications and Huawei. Fig. 7 illustrates a categorization of different research areas with respect to elastic optical networking. Usually, one of the main parameter of interest is the performance on the network level, taking into account aspects as bit rate, transparent reach, and spectral efficiency. On the network level, on the other hand, it is important to examine the overall utilized spectrum, the blocking probability, the utilized interfaces, and the availability. In Fig. 7, a categorization of approaches for network optimization of elastic optical networks is also presented, based on the design scope, the application scope, and the selected methodology.

Novel resource allocation algorithms are required to be developed, as the conventional routing and wavelength assignment (RWA) algorithms of traditional WDM networks can no longer be applied. Instead of assigning a certain wavelength to each connection, a

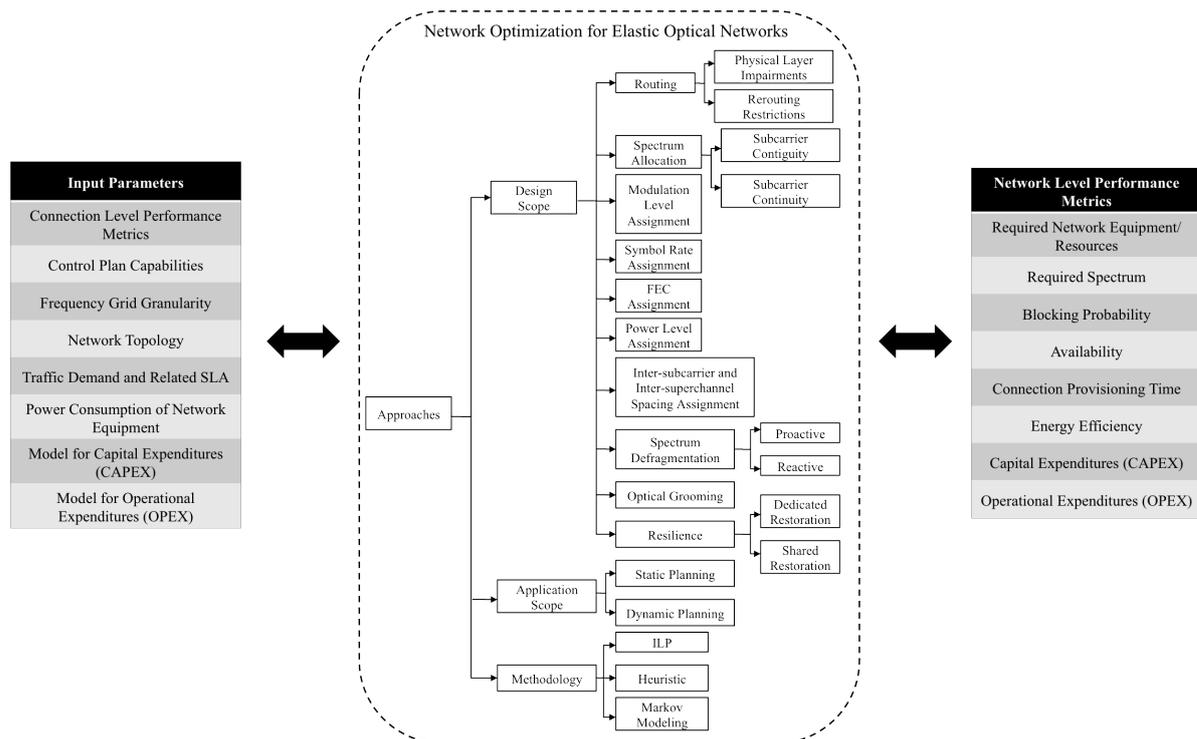


Figure 7: Network optimization for Elastic Optical Networks (adapted from Tomkos *et al.*, 2014).

number of contiguous sub-carrier slots are now to be assigned. Sliceable flexible transceivers can be used in principle to “relax” the *spectrum contiguity constraint*. However, the introduction of such devices might require a more “sophisticated” control manner of the overall network, therefore, increasing the complexity and cost. Additionally, the continuity of these sub-carrier slots should be guaranteed in a similar manner as wavelength continuity constraints are imposed. This leads to the development of routing and spectrum allocation algorithms (RSA) [15, 17, 19, 71]. Moreover, as additional degrees of freedom are allowed by elastic optical networks, new relevant constraints are required to be considered. For example as the modulation level can be selected on a connection basis, constraints tying it to the required bit rate of the traffic demand as well as to the achieved transparent reach are necessary. To this end routing, modulation level and spectrum allocation algorithms (RMLSA) have been recently proposed [14, 16]. Physical layer impairments can be considered in the planning procedure. Note that in this case the maximum transmission distance is a commonly used metric. Additionally, restrictions can be imposed on the manner in which connections are allowed to be re-routed. Re-routing may be desired for example in order to avoid blocking of new connection requests.

The discussed algorithms can address the off-line network planning phase, or they can be applied to dynamically provision connection requests (see Subsection 3.2.2). As connections are dynamically established and released, the issue of bandwidth fragmentation arises (see Section 3.3) – leading to increased blocking probabilities. Thus, the development of spectrum defragmentation algorithms is required. However, optical devices,

which achieve such defragmentation cycles can be quite expensive and even more importantly, defragmentation might interfere with existing traffic. Efficient spectrum allocation schemes, discussed in the following, Subsection 3.3.1, can be used in order to minimize blocking of new connection requests in the context of spectrum fragmentation.

Another categorization involves the method used to conduct the planning. Mathematical optimization methods, such as integer linear programming (ILP) and Markov modeling, can be applied. Heuristic approaches can also be applied – especially if the computational complexity is a restricting factor. In this case, optimality can be sacrificed in order to reduce the computation time. This leads to heuristic methods being the method of choice for dynamic planning. Note that the discussed network planning approaches can be applied with different optimization objectives, such as spectrum savings and blocking performance. However, in terms of blocking performance only few works in the literature had considered Markov-based models for statistical analysis of dynamic resource provisioning in elastic optical networks and, so far, did not provide any in-depth investigation of blocking events and their relation to spectrum fragmentation as presented in this thesis.

3.2 Dynamic Resource Provisioning in EONs: Definition and Complexity

There has been a considerable increase in the range of transmission rates and speed between large and small bandwidth demands that current optical networks are required to provide. Requirements for dissimilar data rates may arise from the geographic distribution of traffic or the variety of services provided. For instance, in some cases high-bit rate traffic (e.g. 100 Gb/s) may be needed for data-center interconnection, while other users may require transport connections of only 100 Mb/s, 1Gb/s and 10 Gb/s. Thus, based on this broad range of traffic granularities, it is desired that optical nodes allocate resources in a flexible and efficient manner to efficiently support high-speed channels (beyond 100G), lower speed channels (e.g. 40 Gb/s, 10Gb/s) and sub-wavelength channels (e.g. hundreds of Mb/s). The desired elastic right-size bandwidth allocation in EONs is achieved with the aid of multi-carrier solutions such as CO-OFDM technology as well as N-WDM, which have set the stage for envisioning fully elastic optical networking. As mentioned in 2.2.2, such multi-carriers solutions technologies enable the useful bandwidth of an optical fiber being discretized and divided into multiple *optical frequency slots* (sub-carriers). Therefore, being the width of a single optical sub-carrier can be much smaller than the width of a channel employed in a fixed-size grid scenario, such as the one defined by the ITU-T in [9]. Based on these assumptions, and considering that the bit rate requested by a connection can be converted into particular spectrum bandwidth needs, each demand

imposed by connection request can be understood as a requested numbers of frequency slots between a source and a destination node.

Elastic bandwidth provisioning enables channels to share the formerly fixed spectral resources, thus improving network efficiency. As show in Fig. 8, a portion of spectrum may be shared in the frequency domain among several services/applications that use similar or different transmission formats, e.g., number of sub-carriers, modulation format as well as bit rate. All the services/applications can coexist as long as they do not interfere with each other. In other words, for a given traffic demand, the request can be translated into a number of optical sub-carriers, and accommodated through the establishment of the corresponding spectrum path. To form the spectrum path for a connection using multiple optical sub-carriers, elastic optical networks may comprise bandwidth variable (BV) transponders at the network edge and bandwidth variable optical cross connects (OXC)s in the network core, which can be built based on the continuous bandwidth variable wavelength selective switch (WSS) [72]. Note that two spectrum paths that share one or more common fiber links, have to be separated in frequency domain to enable the optical signal filtering, i.e., two set of sub-carriers within the two spectrum paths have to be isolated by a *guard frequency band* (guardband). The size of the guardband, however, is not trivial and may be in the order of none, one or multiple sub-carrier(s) [73].

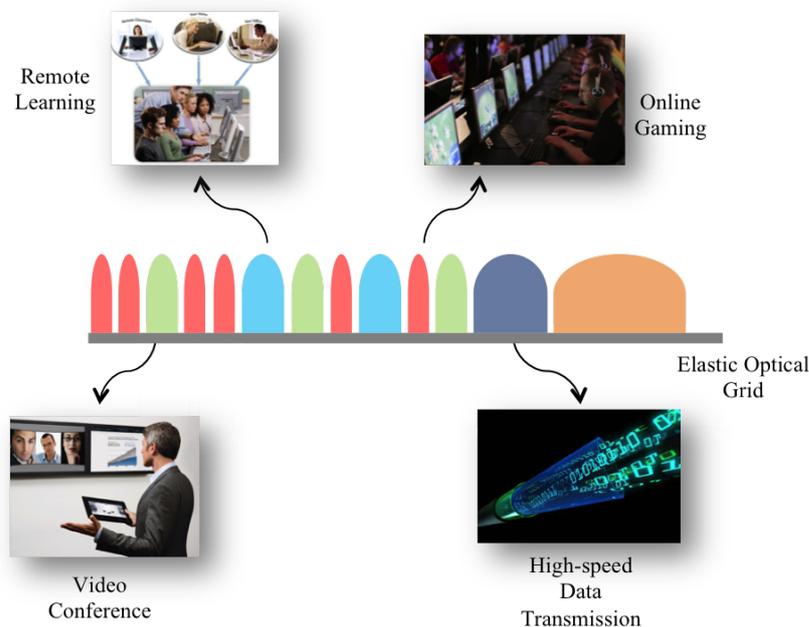


Figure 8: Elastic resource allocation used to carry different services/applications with custom bandwidth allocation.

In the literature, the study in [72] raised the challenges for the future optical networks, while exploring the possibility and feasibility of adopting EONs for next-decade networks. The concept of routing and spectrum allocation was introduced in [72] for the first time, and later studied in [15, 71]. In the following, we formally define the routing and spectrum

allocation (RSA) problem, focusing more on dynamic provisioning of spectral resources and its main issues related to the network level.

3.2.1 Routing and Spectrum Allocation (RSA) Problem

To achieve the elastic and fine-granular bandwidth allocation in elastic optical networks, a similar process as the RWA in the traditional fixed-grid WDM networks, the so called RSA has to be employed. In specific, the RSA process routes and allocates spectrum resources to form the spectrum path, which is an all-optical trail established between the source and link nodes by using one or multiple consecutive optical sub-carriers. Similar to the lightpath in wavelength routed networks (WRN), the spectrum path has to ensure the continuous availability of the allocated sub-carriers along its routing path (*spectrum continuity constraint*). However, the RSA problem is different from and more challenging than the traditional RWA problem due to the following factors. As we pointed before, multi-carrier solutions, such as CO-OFDM, require that for a given spectrum path, the allocated frequency slots have to be consecutive in spectrum domain to be effectively modulated (*spectrum contiguity constraint*). Although the slots of the same spectrum path can be consecutive and overlapping in the spectrum domain, in two spectrum paths have to be separated in the spectrum domain by guardbands when these two spectrum paths share one or more common fiber links. These guard frequencies are used to facilitate the physical frequency filtering. In addition to that, unlike the fixed-grid WDM network where guardband frequencies are pre-allocated and fixed, the guardbands in EON can be any of the sub-carriers and are determined in the process of spectrum paths establishment.

Therefore, the solution for the RWA problem of WRNs cannot be directly applied to elastic optical networks. Similarly, the RSA problem is also different from the routing and wavelength band switching (WBS) networks, where the major goal is to reduce the number of ports in the network. In WBS networks, a number of wavelengths are grouped into a common optical tunnel, namely waveband, and switched as a single entity whenever possible [74]. Conceptually, grouping wavelengths is similar to the allocation of adjacent slots for a given spectrum path in EONs. However, different from the consecutive sub-carriers of a spectrum path, the grouped wavelengths can be from various node-pairs sharing at least one common fiber [74]. The wavelengths within a band are not necessarily adjacent [75], and grouping wavelengths is primarily for the sake of port savings. In contrast, the EON has to ensure the consecutiveness of sub-carriers for effective modulation [4, 73].

One example of routing the spectrum paths using the OXC node in an elastic OFDM-based optical network is shown in Fig. 9, where Fig. 9(a) is a star network with two directional fibers per link and guardband equal to one optical sub-carrier. The BV OXCs in Fig. 9(c) are arranged with a *broadcast-and-select* configuration. The local traffic can be added and dropped through the connection to the OFDM transmitter and receiver,

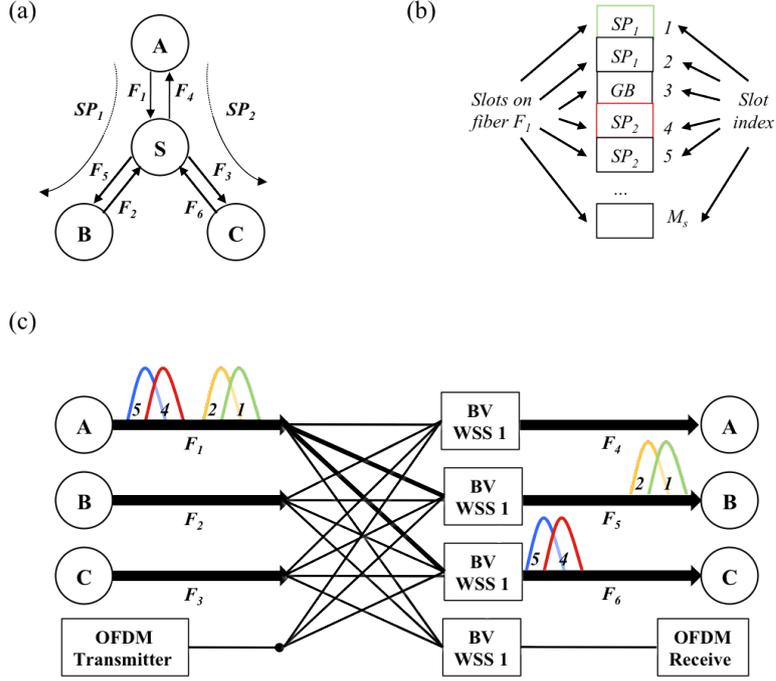


Figure 9: Example of bandwidth selective WXC in the RSA.

respectively. In Fig. 9(a), there is a spectrum path SP_1 of two sub-carriers from A to B , and there is another spectrum path SP_2 of two sub-carriers from A to C . Fig. 9(b) shows the spectrum allocation on fiber F_1 for SP_1 and SP_2 . As shown in Fig. 9(b), each optical sub-carrier on the fiber has an index. The sub-carriers with index 1 and 2 are assigned to SP_1 which requires two adjacent slots. The sub-carriers with index 4 and 5 are assigned to SP_2 . Note that the sub-carriers within SP_1/SP_2 are consecutive and no guardband is needed within SP_1/SP_2 . On the other hand, the sub-carrier with index 3 is assigned as the guardband (GB) between SP_1 and SP_2 since they are overlapping on fiber F_1 . As a result, to accommodate SP_1 and SP_2 , fiber F_1 requires five sub-carriers. Figure 9(c) shows the switching configuration at node S , where the traffic from A to S (through fiber F_1) is sent to BV OXCs 2 and 3 to filter out to the node B or C . Clearly, the required number of slots on fiber F_1 depends on the employed sub-carrier with the maximum index (M_S). Hence, if there are no other traffic demands in Fig. 9(b), M_S of the network will be 5. Fig. 9(c) shows the switching configuration at node S , where the traffic from A to S (through fiber F_1) is sent to BV-WSSs 2 and 3 to filter out to the node B or C . In the following, we formally define the routing and spectrum allocation (RSA) problem in the case with off-line (static) and on-line (dynamic) approaches.

3.2.2 Off-line and On-line Approaches

Recently, there are an increasing number of research works investigating solutions to the RSA problem of elastic optical networks under both static and dynamic traffic scenarios, including considerations of the distance-adaptive modulation technology. The static

solution of RSA deals with the routing and resource allocation during the network planning stage, where an a-priori traffic matrix is given in terms of capacity needed, and the routing and spectrum assignment operations are performed off-line. The capacity requirement of a connection is transformed to a number of sub-carriers (slots), based on the capacity of each sub-carrier. Mathematical optimization methods, such as integer linear programming (ILP) have been proposed to find the optimum solution through a combined routing and spectrum allocation [13, 15, 71]. The objective in this case consists in minimizing the utilized spectrum, considering the *spectrum continuity constraint*, along the path, and the *spectrum contiguity constraint* for each connection requests. The advantages of these methods is that they offer a measure of the optimality of the solution they provide. However, they may lead to increased computational complexity. On the other hand, dynamic network planning approaches are applied to dynamically provision connection requests [17]. Heuristic approaches offer an alternative, especially if the computational complexity is a restricting factor, often making them the method of choice for dynamic planning. However, it is often that, in this case, optimality is sacrificed in order to reduce the computational time.

In networks with non-uniform traffic demands, the solution of RSA is applied to dynamically provision connection requests, and it turns even more challenge since the available spectrum can be highly fragmented, due to the non-uniform spectrum usage. The heterogeneous bandwidth allocation leads to one of the most important problems in EON with dynamic traffic conditions, so called *spectrum fragmentation*. Spectrum fragmentation can highly penalize the overall blocking performance in the network, especially for connection requests demanding high-data rates (i.e., a significant number of continuous slots), as illustrated in Fig. 10. In Fig. 10(a), a certain number of connections, with different bandwidth requirements, is established over a given network link. After some time, in Fig. 10(b), one of these connections is released, thus freeing a portion of the spectrum in the optical link. Finally, in Fig. 10(c), a new high-data rate connection request arrives at the network and should be allocated on the considered optical link. Even though the total spectrum available on the optical link would be enough to allocate the new connection, such spectrum is fragmented into smaller portions than the contiguous spectrum requested by the incoming connection, which eventually causes its blocking.

To mitigate the spectrum fragmentation effects, spectrum defragmentation strategies have been proposed in the literature [76, 77]. Essentially, the idea behind these strategies is to properly rearrange active connections in the network, so as to free as much contiguous spectrum as possible to be used by future connection requests. However, the disruption of active traffic caused by reallocation is not admissible for certain classes of service, being hitless spectrum defragmentation of paramount importance in such a context. In addition to that, defragmentation techniques, add an extra complexity to the control plane in order to perform and manage all the re-allocation properly. On the other hand,

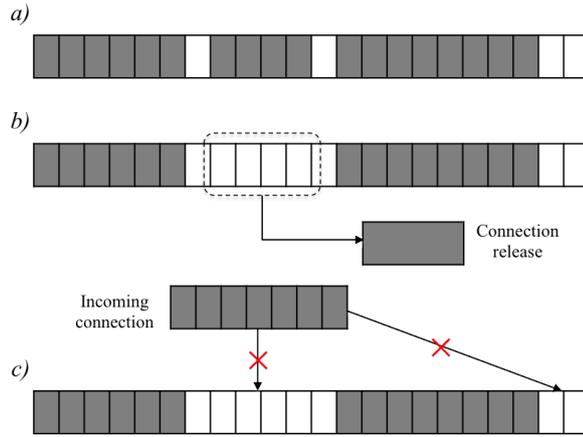


Figure 10: Example of spectrum fragmentation in a dynamic scenario.

equipping the nodes with spectrum conversion capabilities would allow the reduce the blocking probability of some demands through using not necessarily the same spectrum in all the links. However, spectrum conversion devices are expensive and introduce some delay in the demands (due the conversion) that may not be admissible for certain classes of service. Therefore, spectrum allocation (SA) schemes to minimize fragmentation in the optical spectrum have been proposed.

3.3 Spectrum Fragmentation Problem

Spectrum fragmentation is inevitable in networks with non-uniform dynamic bandwidth allocation. The process of adding and removing connections may lead to free spectrum, which is scattered across the considered frequency range. This in turn results in many small non-contiguous frequency bands that cannot be used to satisfy requests of larger bandwidth. Hence, the spectrum efficiency gained by flexibility in the bandwidth allocation may be reduced [76].

The spectrum fragmentation in EON is very similar to the fragmentation in computer memory and storage systems, a problem, which has already been studied extensively [78]. Thus, it should be possible to utilize some measures from this field in order to quantify the spectrum fragmentation in EON. Following this intuition, we propose a way to analyze the relation between blocking probability (BP), spectrum fragmentation and network load, offering a novel way to forecast how much BP will change with spectral fragmentation, considering different SA strategies. We present metrics to measure bandwidth fragmentation, considering four different spectrum allocation policies, which are defined hereafter in the next subsection.

3.3.1 Spectrum Allocation (SA) Policies

In Computer Science, a memory allocator is defined as an on-line algorithm that must respond immediately to a request in a strict sequence [78]. The main objective of an allocator is to minimize wasted space without undue time cost. Usually, allocators are categorized by the mechanism to trace, which blocks are free, and to merging adjacent free blocks into free larger ones.

Obviously, memory allocators fulfill the same principal purpose as spectrum allocation entities in EON. In both environments it is important to have suitable strategies in order to reduce the waste of resources. In our study, we adapt the concept of such allocators. Based on this, in [35] we proposed a novel spectrum allocation policy called *Exact-Fit*; and compared it with the existing ones in terms of average BP and fragmentation based on an analytical modeling approach for flexible spectrum allocation under dynamic traffic conditions.

The following existing dynamic SA schemes are compared in this thesis:

- *First-Fit (FF)*: the *First-Fit* policy places the request in the first available frequency band large enough to meet the demand;
- *Smallest-Fit (SF)*: this policy allocates the smallest free block, hence filling-up gaps in order to reduce fragmentation;
- *Random-Fit (RND)*: the *Random-SA* policy allocates incoming requests in any available block large enough to satisfy the requested bandwidth and is considered for benchmarking purposes.

The proposed spectrum allocation policy works in the following way:

- *Exact-Fit (EF)* [35]: Starting from the beginning of the frequency channel, *EF* searches for an available block with the number of slots equal to the one requested by the connection. If there is a block to match the exact size of requested resources, the algorithm allocates that spectrum. Otherwise, the spectrum is allocated in the largest free block.

3.3.2 Fragmentation Ratio Calculation

Different alternatives to quantify the spectrum fragmentation in EON have been proposed. In our previous work [35], we analyzed two measures of fragmentation. The first one is based on the fragmentation calculation introduced in [78] and also used in [29], which takes into account the number of slots in the largest free block over the total number of

available slots in the system. Equation (3.1) expresses this measure, which is valid under the assumption that there is always a free slot available in the spectrum.

$$Frag = 1 - \frac{\#slots \text{ in largest free block}}{\text{total number of free slots}} \quad (3.1)$$

The second measure introduced in [35] relates the number of slots required to serve a certain connection demand and the capability to meet such a connection request. The motivation reflects the fact that the impact of fragmentation is dependent on the bandwidth demand of a connection request, e.g., it looks more severe when a connection requires more spectrum resources. Such definition allows directly analyzing fragmentation as a function of the number of contiguous slots required by a connection request. This measure is expressed in (3.2).

$$Frag(n_{req}) = 1 - \frac{n_{req} \times \#blocks(n_{req})}{\text{total number of free slots}} \quad (3.2)$$

where n_{req} corresponds to the number of slots required by a certain connection request and $\#blocks(n_{req})$ represents the number of non-overlapping blocks of n_{req} slots, i.e., how many times such connection request can be simultaneously served in a certain network state. The denominator corresponds to the total number of remaining free slots. Equation (3.2) is valid under the assumption that there is always a free slot in the system. The value of fragmentation defined according to (3.2) changes depending on the size of the connection request, in terms of contiguous slots required – in the following also referred to as the type of a request.

Both aforementioned metrics capture the essentials of how much free spectrum is not used efficiently due to the fragmentation. It is intuitive that, given a certain spectrum occupancy, every SA policy should result in different fragmentation values. That is, for a given network state and depending on the connection requests, different SA policies might cause different fragmentation. By analyzing our numerical results, we found that the measure of fragmentation according to (3.1) does not differentiate properly the details of fragmentation resulting from using different SA policies, because it only evaluates the largest free blocks. It does not consider any distribution of other smaller free blocks or number of their appearance.

Therefore we use a metric of fragmentation introduced in [79] that can actually be seen as an extension of the measure in [78].

$$Frag(n_{req}) = \frac{Frag_{int}(n_{req}) + Frag_{ext}(n_{req})}{\text{total number of free slots}} \quad (3.3)$$

where

$$Frag_{int}(n_{req}) = \begin{cases} m = \tilde{m} - n_{req}; & m < \min(n_{req}) \\ 0; & \text{otherwise} \end{cases} \quad (3.4)$$

$$Frag_{ext}(n_{req}) = \sum_{i: (\#slots \text{ of block } i) < n_{req}} \#slots \text{ of block } i \quad (3.5)$$

$Frag_{int}$ (in [79] referred to as *internal* fragmentation) considers the number of remaining slots in the idle spectrum interval of size \tilde{m} slots, which a request of n_{req} slots would be allocated in, and cannot be used for any future request no matter the amount of resources demanded. The remaining slots in such empty block are even smaller than the smallest request size and therefore are considered wasted. $Frag_{ext}$ (in [79] referred to as *external* fragmentation) accounts for the idle spectrum intervals that cannot be used for the actual request of n_{req} slots and are considered holes in between the otherwise allocated blocks. To calculate $Frag_{ext}$ in (3.5), the sum is over the number of slots belonging to all the unallocated blocks of slots, whose sizes are smaller than a required number of slots n_{req} for the incoming request.

To give an example, suppose an optical spectrum with 31 OFDM sub-carriers slots where a slot granularity is 12.5 GHz. Let us consider the spectrum allocation for two types of connections, which require 37.5 GHz and 25 GHz transmission bandwidth, respectively, plus additional n_G guardband slots (here $n_G = 1$), on either side, to separate the adjacent connections [13], as illustrated in Fig. 11. We note that one guardband slot shared between two connections could be considered sufficient, if guardband allocation is coordinated between adjacent connections. However, since it does not contribute significantly to the principal problem of fragmentation, we disregard this aspect in the following explanations and generally assume $n_G = 1$ on either side of the spectrum requested by the connections.

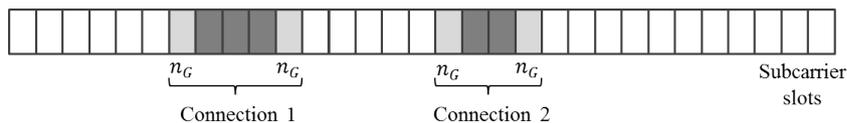


Figure 11: Example of spectrum configuration.

Suppose that an additional connection arrives to the system. The request is assumed to require two slots and two guardband slots, i.e., in total four slots are needed to serve this connection request. These four slots can be allocated in different ways, depending on the SA policy applied. Possible “spectrum allocation configurations” resulting from different spectrum allocation strategies are illustrated in Fig. 12.

Using the fragmentation rate calculation expressed in (3.3) and considering that the smallest bandwidth required in the system is equal to two contiguous slots (plus two guardband slots) results in $(a) = 0,11\%$; $(b) = 5\%$; $(c) = 0\%$; $(d) = 16\%$ for FF , SF , EF and RND , respectively.

Table 1 compares the values of fragmentation in % of the sample spectrum configuration as illustrated in Fig. 10 for spectrum fragmentation according to (3.1) and the approach according to (3.3). It can be seen that values of fragmentation according to [78],

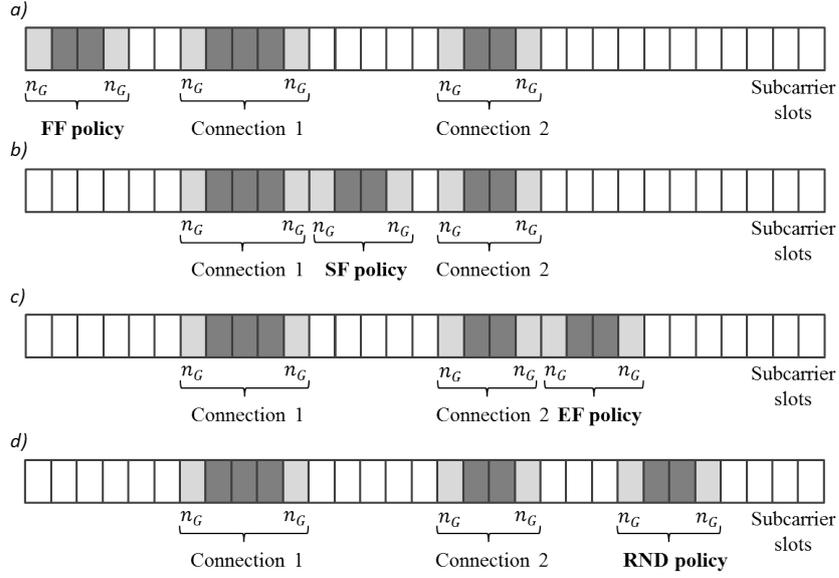


Figure 12: Possible spectrum configuration resulting from (a) *FF*, (b) *SF*, (c) *EF* and (d) *RND* SA policies.

i.e., metric (3.1), are the same in case of *First-Fit* and *Smallest-Fit* SA policies, even though the spectrum utilization already looks different in our small example. The reason for this is that for metric (3.1) the largest free spectrum interval is mostly unaffected by *FF* and *SF*. For *FF*, the largest free spectrum interval is considered for allocation only, if this block is, by coincidence, the first block large enough, which literally is not what the policy of *FF* enforces. For *SF*, the largest free spectrum interval is almost impossible to be considered, again due to the nature of the *smallest-fit* SA policy. A similar reasoning as for *SF* holds for the definition of *EF*. Metric (3.3), on the opposite, does directly take into account the remaining in the empty block considered for allocation through *Frag_{int}*. It does therefore reflect the change in spectrum occupancy, no matter where a request is allocated. In addition, *Frag_{ext}* also accounts better on the actual type of allocation request than by examining empty blocks that are least likely relevant.

Table 1: Comparison of fragmentation metrics resulting from different SA policies

	Fragmentation According to (3.1)	Fragmentation According to (3.3)
<i>FF</i>	38%	11%
<i>SF</i>	38%	5%
<i>EF</i>	61%	0%
<i>RND</i>	66%	16%

The sample spectrum configurations as depicted in Fig. 12 are considered very simple

and not too scattered. However, they can serve as examples for not obviously different spectrum configurations obtained by various SA strategies and, consequently, point out the need for proper fragmentation metrics able to distinctly classify the performance of the SA strategies. We have evaluated the fragmentation metrics expressed in Eq. (3.1) and (3.3) using the elastic SA framework, proposed in Chapter 4, in order to identify the most relevant one in terms of reflecting the relation between fragmentation level and the link-performance, e.g., blocking probability. By using the metric in (3.3), it was possible to identify some very small differences in spectrum configurations, which were not recognizable when we considered the measure presented in (3.1).

In this thesis, we use (3.2) and (3.3) to quantify the level of fragmentation per type of connection request and the average system fragmentation, respectively.

4 ELASTIC SA FRAMEWORK BASED ON MARKOV MODELING

In this Chapter, we introduce the planning framework used in this thesis. We first explain the overall process of Markov modeling in Section 4.1. In Subsection 4.1.1, we explain continuous time Markov chains (CTMCs), and review some principal methods to efficiently compute the steady-state probabilities in Subsection 4.1.2. In this thesis, we consider Markov modeling of elastic optical links, which was first presented in our conference paper [35]. Then, as an extension of our conference paper, in our journal contribution [36], we propose an analytical framework, based on Markov modeling, to calculate steady-state probabilities in order to investigate the inherent reasons for blocking events. These contributions are covered in Section 4.2. We begin with an introduction of the space-state generation methodology in Subsection 4.2.1, and move on to the transition-rate matrix generation and then, show how the steady-state probabilities are obtained in Subsection 4.2.2. Finally, in Section 4.3, we present the new definitions for blocking events that are used to statistically analyze the performance of different SA schemes studied in this thesis.

4.1 Markov Modeling

Computer and communication systems are ubiquitous in all spheres of our life. Discrete-state models have proved to be a valuable tool in the analysis of these computer systems and communication networks. Modeling of such systems involves the description of the system's behavior by the set of different states the system may occupy, and identifying the transition relation among the various states of the system. Uncertainty is an inherent feature of real-life systems and, to take account of such behavior, probability distributions are associated with the possible events (transitions) in each state, so that the model implicitly defines a stochastic process. If the probability distributions are restricted to be either geometric or exponential, the stochastic process can be modeled as a discrete time (DTMC) or a continuous time Markov chain (CTMC) respectively. A Markov decision process (MDP) admits a number of discrete probability distributions enabled in a state which are chosen non-deterministically by the environment. In this thesis we concentrate on continuous time Markov chains.

The overall process of the state based analytical modeling for CTMCs involves the specification of the system, the generation of the state-space, and the numerical computation of all performance measures of interest. Specification of a system at the level of a Markov chain, however, is difficult and error-prone. Consequently, a wide range of high-level

formalisms have been developed to specify the system under study. See, for example, [80] for a survey of model representations. Once a system is specified using some suitable high-level formalism, the entire state-space needs to be generated from this specification. In the following, we define continuous time Markov chain and its main properties.

4.1.1 Continuous Time Markov Chains (CTMC)

As aforementioned, a system may be represented as a stochastic process by describing the set of different states the system may occupy and by identifying the transitions which can occur between the various states of the system. A stochastic process is a family of random variables $\{X(t), t \in T\}$ indexed by t , usually a time parameter. The random variable $X(t)$ denotes an observation of the system at time instant t . A stochastic process with discrete time indices, for example, $T = \{0, 1, 2, \dots\}$, is called a *discrete time parameter* stochastic process; if time is continuous, e.g. $T = \{0 \leq t \leq +\infty\}$, it is a continuous time parameter stochastic process. The values that random variable $X(t)$ can take are called states, and the set of all possible values constitutes the state-space of the process. If the values assumed by the variable $X(t)$ are discrete, it forms a discrete state-space.

A Markov process is a stochastic process which satisfies the Markov property, i.e., for all positive integers k , any sequence of time instances $t_0 < t_1 < \dots < t_k$ and states x_0, \dots, x_k :

$$P[X(t_k) \leq x_k | X(t_k - 1) = x_{k-1}, \dots, X(t_0) = x_0] = P[X(t_k) \leq x_k | X(t_{k-1}) = x_{k-1}]. \quad (4.1)$$

The Markov property formulated above, sometimes known as the *memoryless property*, implies that the state in which the system finds itself at time t_k depends only on the state of the system at time t_{k-1} , while the state occupied by the system at any previous time instances (i.e. t_0, t_1, \dots, t_{k-2}) is completely irrelevant. Consequently, the future state of the system is also independent of the time spent so far in the current state of the system. Note that it is still possible for the transitions to depend on the actual time at which they occur. In this thesis, however, we only consider the homogeneous case where the transitions are independent of time.

A Markov process with a discrete state space is referred to as a *Markov chain*. Accordingly, a Markov chain with a continuous time parameter (t) is called a *continuous time Markov chain* (CTMC). A CTMC is a continuous time, discrete-state stochastic process, i.e. it is a Markov chain which can change state at any time instant. Mathematically, a CTMC is a stochastic process $\{X(t), t \geq 0\}$ which satisfies the Markov property given by

equation (4.2), which in this case (continuous state) can be formulated as:

$$P[X(t_k) = x_k | X(t_{k-1}) = x_{k-1}, \dots, X(t_0) = x_0] = P[X(t_k) = x_k | X(t_{k-1}) = x_{k-1}] \quad (4.2)$$

for all positive integers k , any sequence of time instances $t_0 < t_1 < \dots < t_k$ and x_0, \dots, x_k . The only continuous probability distribution, which satisfies the Markov property is the exponential distribution.

A CTMC may be represented by a set of states S , and the *transition-rate matrix* $R : S \times S \rightarrow \mathbb{R}_{\geq 0}$. A transition from state i to state j is only possible if the matrix entry $r_{ij} > 0$. The matrix coefficients determine transition probabilities and state *sojourn times* (or *holding times*). Given the *leaving rate* of state i , $E(i) = \sum_{j \in S, j \neq i} r_{ij}$, the mean holding time for state i is $1/E(i)$, and the probability of making a transition out of state i within t time units is $1 - e^{-E(i)t}$. When a transition from state i does occur, the probability that it goes to state j is $r_{ij}/E(i)$.

An infinitesimal generator matrix Q may be associated to a CTMC by setting the off-diagonal entries of the matrix Q with $q_{ij} = r_{ij}$, and the diagonal entries with $q_{ii} = -E(i)$. The matrix Q (or R) is usually sparse; further details about the properties of these matrices can be found in [81]. In general, when analyzing CTMCs, the performance measure of interest corresponds to either the probability of being in a certain state at a certain time (*transient*) or the long-run (*steady-state*) probability of being in a state. Transient state probabilities can be determined by solving a system of ordinary differential equations. The computation of steady-state probabilities involves the solution of a system of linear equations. We will focus in this thesis on the steady-state solution of a CTMC. See the next section for more details.

Finally, we define the notion of reachability. We say that there exists a transition from state i to state j , if $q_{ij} > 0$. A state j in a CTMC is reachable from another state i in the CTMC, if there exists a finite sequence of transitions in the model from the state i to the state j . The set of all states which can be reached from the initial state is called the set of reachable or possible states of the model.

4.1.2 Numerical Methods

Let be $Q \in \mathbb{R}^{n \times n}$ the infinitesimal generator matrix of a continuous time Markov chain with n states, and

$$\pi(t) = [\pi_0(t), \pi_1(t), \dots, \pi_{n-1}(t)] \quad (4.3)$$

the transient state probability row vector, where $\pi(t)$ denotes the probability of the CTMC being in state i at time t . The transient behavior of the CTMC is described by the

Chapman-Kolmogorov differential equation:

$$\frac{d\pi(t)}{dt} = \pi(t)Q. \quad (4.4)$$

To compute this, the initial probability distribution of the CTMC, $\pi(0)$, is also required. We concentrate on the steady-state behavior of a CTMC, which is obtained by solving the system of equations:

$$\pi \times Q = 0 \quad (4.5)$$

under the constraint

$$\sum_{i=0}^{n-1} \pi_i = 1, \quad (4.6)$$

where

$$\pi = \lim_{t \rightarrow \infty} \pi(t) \quad (4.7)$$

is the *steady-state probability* vector. A sufficient condition for the unique solution of (4.5) is that the CTMC is finite and irreducible. A CTMC is *irreducible* if every state can be reached from every other state [82], for example. Equation (4.5) can be reformulated as $Q^T \pi^T = 0$, and well-known methods for systems of linear equations of the form $Ax = b$ can be used to simultaneously solve for (4.5) and (4.6).

The numerical solution methods for linear systems of the form $Ax = b$ are broadly classified into two categories: *direct methods*, such as *Gaussian elimination*, *LU factorization*, etc; and *iterative methods*. Direct methods obtain the exact solution in finitely many operations and are often preferred to iterative methods in real applications because of their robustness and predictable behavior. However, as the size of the systems to be solved increases, they often become almost impractical due to the phenomenon known as *fill-in*. The fill-in of a sparse matrix is a result of those entries which change from an initial value of zero to a nonzero value during the factorization phase, e.g. when a row of a sparse matrix is subtracted from another row, some of the zero entries in the latter row may become nonzero. Such modifications to the matrix mean that the data structure employed to store the sparse matrix must be updated during the execution of the algorithm.

Iterative methods, on the other hand, do not modify matrix A ; rather, they involve the matrix only in the context of matrix-vector product (MVP) operations. The term “iterative methods” refers to a wide range of techniques that use successive approximations to obtain more accurate solutions to a linear system at each step [83]. Beginning with a given approximate solution, these methods modify the components of the approximation, until convergence is achieved. They do not guarantee a solution for all systems of equations. However, when they do yield a solution, they are usually less expensive than direct methods. They can be further classified into *stationary methods* like *Jacobi* and

ALGORITHM 1: Standard Jacobi Algorithm

```
1: while not converged do
2:    $a = (L + U)\tilde{x}$ 
3:    $x \leftarrow b - a$ 
4:    $x \leftarrow D^{-1}x$ 
5:   Test for convergence (compare  $x$  and  $\tilde{x}$ ) and stop if accuracy sufficient
6:    $\tilde{x} \leftarrow x$ 
7: end while
```

Gauss-Seidel, and *non-stationary* methods such as *Conjugate Gradient*, *Lanczos*, etc. In the following, we will briefly describe some iterative methods to obtain our steady-state probabilities according to (4.5) and (4.6), considering that Q is stochastic and (4.6) can substitute one equation in the system (4.6).

4.1.2.1 Jacobi Method

We now consider the iterative methods for the solution of the system of equations $Ax = b$, where A is square of size n . In the k -th iteration of the Jacobi method, we calculate:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij} x_j^{(k-1)} \right) \quad (4.8)$$

for $0 \leq i, j < n$, assuming indexing starts at 0, where a_{ij} denotes the element in row i and column j of matrix A and the term $x_i^{(k)}$ indicates the i -th element of the k -th iteration vector. The above equation can also be written in matrix notation as:

$$x^{(k)} = -D^{-1}(L + U)x^{(k-1)} + D^{-1}b, \quad (4.9)$$

where D , L and U are a partitioning of A into its diagonal, strictly lower- and upper-triangular parts, respectively. The Jacobi method, can be formulated into an matrix-vector product (MVP) based algorithm, for example, as shown in Algorithm 1.

In Algorithm 1, \tilde{x} represents the old estimate of the solution for x . Line 2 of the algorithm performs the MVP operation, the most complex step. The algorithm requires storage for two iteration vectors (the previous approximation \tilde{x} and the current one x), for the matrix $L + U$ and for the diagonal entries in D . Note that the new approximation of the solution vector is calculated using only the old approximation of the solution. This makes the Jacobi method well suited for parallelization, but means it tends to exhibit slow convergence.

4.1.2.2 Gauss-Seidel Method

The Gauss-Seidel method, which in practice converges faster than the Jacobi method, uses the most recently updated approximation of the solution as soon as available:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij}x_j^{(k)} - \sum_{j>i} a_{ij}x_j^{(k-1)} \right) \quad (4.10)$$

for $0 \leq i, j < n$. The other advantage of the Gauss-Seidel algorithm is that it can be implemented using only one iteration vector. The Gauss-Seidel method can also be expressed in matrix notation:

$$x^{(k)} = -(D + L)^{-1}Ux^{(k-1)} + (D + L)^{-1}b \quad (4.11)$$

or

$$x^{(k)} = D^{-1}(-Lx^{(k)} - Ux^{(k-1)} + b) \quad (4.12)$$

where D , L and U are as described for the Jacobi method above. In practice, the inverse in the first matrix notation would be computed sequentially using forward substitution. In the second form it is visible how updated estimates of elements of x can be immediately re-used, since in the product $Lx^{(k)}$, updated elements of x do not affect themselves, different to the case in (4.9).

4.1.2.3 Successive Over-Relaxation (SOR) Method

An SOR iteration is given by:

$$x_i^{(k)} = \omega \hat{x}_i^{(k)} + (1 - \omega)x_i^{(k-1)}, 0 \leq i < n \quad (4.13)$$

where \hat{x} denotes a Gauss-Seidel iterate, and $\omega \in (0, 2)$ is the relaxation factor. The method is under-relaxed for $0 < \omega < 1$, and is over-relaxed for $\omega > 1$; the choice $\omega = 1$ reduces SOR to Gauss-Seidel method. It is shown in [84] that SOR fails to converge if $\omega \notin (0, 2)$. For good choices of ω , SOR can have considerably better convergence behavior than Gauss-Seidel. However, a-priori computation of an optimal value for ω is not feasible.

4.1.2.4 Sparse Equations and Least Squares (LSQR) Method

LSQR is an implementation of a *Conjugate Gradient* type method for the solution of sparse linear equations and sparse least-squares problems:

$$\begin{aligned} &\text{Solve } Ax = b \\ &\text{or minimize } \| Ax - b \|^2 \\ &\text{or minimize } \| Ax - b \|^2 + \lambda^2 \| x \|^2, \end{aligned}$$

where the matrix A may be square or rectangular (over-determined or under-determined), and may have any rank. It is represented by a routine for computing $A\nu$ and $A^T u$ for given vectors ν and u . The scalar λ is a damping parameter. If $\lambda > 0$, the solution is “regularized” in the sense that a unique solution always exists, and $\|x\|$ is bounded.

The method is based on the bi-diagonalization procedure of Golub and Kahan. It is analytically equivalent to the standard method of Conjugate-Gradient; however, it possesses more favorable numerical properties, especially if A is ill-conditioned. The reader is referred to [85] for specific details and mathematical proof of the LSQR method.

4.2 Elastic SA Framework

The status of slot allocation over time in an elastic optical link can be modeled as a Markov chain, which will describe the stochastic process corresponding to the spectrum occupancy in the system. It will allow obtaining the probabilities of the system to be in given states of spectrum usage.

We assume that the spectrum of an elastic optical link (EOL) is divided in N slots and we consider a set of C different types $c = 1 \dots C$ of possibly incoming connection requests, whereas a type c request demands n_c contiguous slots. Each slot is considered to have a status of being either unused or allocated/busy due to a successfully granted connection request. The status combination for all slots is a sufficient representation of the overall spectrum utilization of the elastic optical link, in the following also referred to as the state of the EOL.

The states of the EOL and the transition relations between them can be, under certain assumptions, modeled as a Markov chain, where the steady-state probabilities allow quantifying the blocking probability and fragmentation on the link under dynamic traffic conditions. The steady-state probabilities can be derived by the elastic SA framework, presented in a flow diagram in Fig. 13. The components of this framework are described hereafter in the next subsections.

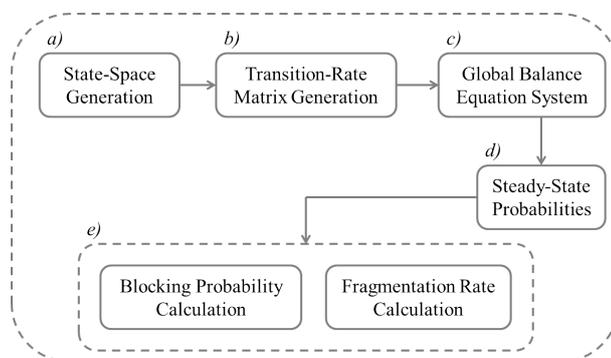


Figure 13: Flow diagram for the Elastic SA framework.

4.2.1 State-Space Generation

Each state in the Markov chain is represented by an integer array of length N , where the i -th element contains the status of slot i . The status of an available slot is indicated by 0; the status of allocated slots is represented by 1 or 2, depending on whether they are used for transmission or as guardband, respectively. The different status indication of guard slots from otherwise allocated slots allows to uniquely identify individual connections. This is essential to trace transitions in the Markov chain resulting from termination of formerly allocated connections.

The state-space is generated by the iterative procedure presented in Algorithm 2. Assume an already generated subset $S = \{s_i\}$ of all possible states of the EOL. For each state s_i , we consider all possible types $c = 1 \dots C$ of incoming and leaving connection requests of respective sizes n_c . If a connection request cannot be accepted due to lack of contiguous slots, the state remains unchanged, while if a request of type c can be allocated, a resulting child state, e.g., a_c is derived as shown in Step 4. Note that the set of newly generated states can vary depending on the actual SA policy. Then, each child a_c is checked whether it is already contained in S in Step 5 and appended only if not. The same procedure as described above is applied for the de-allocation of existing connections, i.e., all possibilities to leave state s_i are assumed, resulting in a possible set of child states $\{d_c\}$, $c = 1 \dots C$ in Step 10. Again, each child d_c is appended only if it is not already contained in S in Step 11-12.

By this procedure, S is iteratively expanded. The procedure is repeated for all states s_i in S that are not considered yet. Note that this includes also the states that are appended to S throughout former iterations. In the pseudo-code, $|S|$ denotes at any time the size of S , i.e., the number of current states in S . Steps from 2 to 16 are repeated until the algorithm converges, i.e., when all states in S are analyzed. At this point, S has converged to the set of all possible states in the Markov chain. Thus, if N_{SA} denotes the number of states for a given SA policy, we have $|S| = N_{SA}$. The set S of states is initialized with $S = \{s_1\}$, s_1 being the state with no slots allocated. To be able to track state transitions, each state s_i is annexed with its parent states, i.e., the states that can transition into s_i , as attributes.

The overall complexity in statistical modeling of elastic optical links grows exponentially with the number of states of the Markov chain. In the following, the number of states N_{RND} for *Random*-SA policy is calculated. Let the elements of $W = \{w_c\}$, $c = 1 \dots C$, represent the number of allocated connections of size n_c . Then, the number of used/busy and available/free slots can be expressed as in (4.14) and (4.15), respectively.

$$A(W) = \sum_{c=1}^C n_c w_c \quad (4.14)$$

ALGORITHM 2: State-Space Generation

```
1: Initialize:  $S = [s_1]$ ,  $i = 1$ 
2: repeat
    // Evaluate all possible connection incoming-requests
3:   for  $c = 1 \dots C$  do
    // Base on SA policy P, generate a new link state
    // after an attempt to allocate a request of  $n_c$  slots
4:      $a_c = P(s_i, n_c)$ 
5:     if  $a_c \notin S$  then
6:        $S \leftarrow [S, a_c]$  // Append new state to  $S$ 
7:     end if
8:   end for

    // Evaluate all possible connection leaving-requests
9:   for  $c = 1 \dots C$  do
    // Generate a new link state after de-allocating a
    // connection of  $n_c$  slots
10:     $d_c = D(s_i, n_c)$ 
11:    if  $d_c \notin S$  then
12:       $S \leftarrow [S, d_c]$  // Append new state to  $S$ 
13:    end if
14:  end for
15:   $i \leftarrow i + 1$ 
16: until  $i > |S|$  // All states are analyzed
```

$$F(W) = N - A(W) \quad (4.15)$$

Of course, a certain W can only occur, if the number of allocated slots does not exceed the number N of available ones, i.e., any considerable tuple W must fulfill Inequality 4.16.

$$A(W) \leq N \quad (4.16)$$

The number of states for *Random*-SA policy is equal to the number of possible combinations to allocate blocks of n_c slots into N available slots and can be expressed by

$$N_{RND} = \sum_{w_1=0}^{\infty} \sum_{w_2=0}^{\infty} \dots \sum_{w_C=0}^{\infty} \mathbb{I}(A(W) \leq N) \times J(W). \quad (4.17)$$

$\mathbb{I}(\cdot)$ is an indicator function taking a Boolean argument and defined as

$$\mathbb{I}(b) = \begin{cases} 1; & b \text{ is true} \\ 0; & \text{otherwise} \end{cases} \quad (4.18)$$

$J(W)$ is the number of possible combinations to place w_c , $c = 1 \dots C$, blocks of n_c slots in N available slots and it can be derived as in the following.

There are

$$B(W) = \sum_{c=1}^C w_c$$

allocated connections or contiguous blocks of slots in an elastic optical link. Therefore, there are $B(W)!$ permutations to order such $B(W)$ allocated blocks. As the order of allocated blocks of identical size does not contribute in differentiating the states of an EOL, there are effectively

$$\frac{B(W)!}{w_1! \times w_2! \times \dots \times w_C!}$$

combinations of allocated blocks left to be taken into account.

Considering that $F(W)$ slots, as defined in (4.15), might be unused, there are

$$\binom{F(W) + B(W)}{B(W)}$$

combinations to place the $B(W)$ allocated blocks in different ways.

We therefore obtain

$$J(W) = \frac{B(W)!}{w_1! \times w_2! \times \dots \times w_C!} \times \binom{F(W) + B(W)}{B(W)} \quad (4.19)$$

To eliminate the indicator function $I(\cdot)$ and in order to make the calculation in (4.17) practical, we consider the range of feasibility for each w_c in (4.17) and define

$$M_1 = \left\lfloor \frac{N}{n_1} \right\rfloor$$

$$M_i(w_1, \dots, w_{i-1}) = \left\lfloor \frac{N - \sum_{c=1}^{i-1} w_c n_c}{n_i} \right\rfloor, \text{ for } i = 2 \dots C. \quad (4.20)$$

The maximum number of states for *Random-SA* policy can be expressed as

$$N_{RND} = \sum_{w_1=0}^{M_1} \sum_{w_2=0}^{M_2(w_1)} \dots \sum_{w_C=0}^{M_C(w_1, \dots, w_{C-1})} J(W) \quad (4.21)$$

The maximum number of states for *First-Fit*, *Smallest-Fit* and *Exact-Fit* SA is smaller than for the *Random-SA* policy. Their states are a sub-set of the states for *Random-SA* policy and obtained according to each SA policy's (*FF*, *SF* and *EF*) formation law.

4.2.2 Transition-Rate Matrix Generation and Steady-State Probabilities

Not only the state-space, but also the transition rates in the Markov model of an EOL are dependent on the spectrum allocation policies. For the calculation of steady-state

probabilities, we assume a continuous time Markov chain (CTMC) with exponentially distributed interarrival and holding times.

Let λ_c and μ_c denote the arrival and service rates of connection requests of type c , respectively. We define further

$$A(s_i, c) = \begin{cases} \text{false;} & \text{if a type } c \text{ connection is blocked in } s_i \\ \text{true;} & \text{otherwise} \end{cases} \quad (4.22)$$

$$L(s_i, c) = \begin{cases} \text{true;} & \text{if a type } c \text{ connection is allocated in } s_i \\ \text{false;} & \text{otherwise} \end{cases} \quad (4.23)$$

$$T(s_i, s_j, c) = \begin{cases} \text{true;} & \text{if transition between states } s_i \text{ and } s_j \\ & \text{is possible by either allocation or due to} \\ & \text{a termination of a type } c \text{ connection} \\ \text{false;} & \text{otherwise} \end{cases} \quad (4.24)$$

Let $a(s_i, c)$ denote the number of different ways a connection of type c can be allocated in state s_i ; $a(s_i, c)$ depends on the SA policy and is always equal to 1 for *FF*, *SF* and *EF*. Let $d(s_i, c)$ denote the number of connection of type c that are actually allocated in state s_i . Remember that in Algorithm 2, we annex state s_i with its parent states and therefore $A(s_i, c)$, $L(s_i, c)$ as well as $a(s_i, c)$, $d(s_i, c)$ are easy to derive. $T(s_i, s_j, c)$ is reciprocal in s_i and s_j , meaning that if a transition from s_i to s_j is possible due to an allocation of a type c request, a transition from s_j to s_i is possible due to a de-allocation of a type c request. Let $\pi(s_i)$ denote the steady-state probability of a state s_i in the Markov chain. Then the relation between the arrival and service rates, the allocation and de-allocation request types and the $\pi(s_i)$ for the assumed model is given as

$$\left(\sum_{c=1, A(s_i, c)}^C \lambda_c + \sum_{c=1, L(s_i, c)}^C d(s_i, c) \mu_c \right) \cdot \pi(s_i) = \sum_{j=1, j \neq i}^{N_{SA}} \left(\sum_{c=1, T(s_i, s_j, c)}^C \frac{\lambda_c}{a(s_j, c)} + \sum_{c=1, T(s_i, s_j, c)}^C \mu_c \right) \cdot \pi(s_j), \quad i = 1, \dots, N_{SA} \quad (4.25)$$

Equation (4.25) represents our EOL specific *Global Balance Equation* (GBE) system [82] that describes the probability flux in and out of state s_i on the RHS and LHS, respectively. The scaling of λ_c in the RHS of (4.25) is necessary, since the fact of multiple options to allocate requests of type c in state s_j (resulting in other states than s_i) must not increase the overall leaving rate of state s_i .

In order to explain how we solve (4.25) for $c = 1 \dots N_{SA}$ simultaneously, we rewrite

(4.25) as

$$\pi(s_i) \sum_{j \neq i} q_{i,j} = \sum_{j \neq i} \pi(s_j) q_{j,i}, \quad i = 1, 2, \dots, N_{SA} \quad (4.26)$$

The $q_{j,i}$ are the transition rates from state j to i in the CTMC. In (4.26), we have $q_{j,i} = \lambda_c$, if a transition from state j to i is possible and caused by a request arrival of type c ; if a transition from state j to i is possible and is caused by termination of a connection of type c , we set $q_{j,i} = \mu_c$. On the RHS of (4.26), the $q_{j,i}$ that are related to a connection arrival event, causing a state transition from state j to i , need to be rescaled by $a(s_j, c)$ in case *Random-SA* policy is considered.

For $i = 1 \dots N_{SA}$, (4.26) can be written as

$$Q^T \times \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_{N_{SA}} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.27)$$

where we abbreviated $\pi_i = \pi(s_i)$, Q^T being the transpose of Q and Q is defined as

$$Q = \begin{bmatrix} -\sum_{j \neq 1} q_{j,1} & q_{1,2} & q_{1,3} & \cdots & q_{1,N_{SA}} \\ q_{2,1} & -\sum_{j \neq 2} q_{j,2} & q_{2,3} & \cdots & q_{2,N_{SA}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{N_{SA},1} & q_{N_{SA},2} & q_{N_{SA},3} & \cdots & -\sum_{j \neq N_{SA}} q_{j,N_{SA}} \end{bmatrix} \quad (4.28)$$

Matrix Q is the so called infinitesimal generator or transition-rate matrix, a mathematical representation of the CTMC, which is commonly used to calculate the Markov chain's stationary distribution.

Obviously, the stationary distribution of the Markov chain must fulfill the relation as expressed in (4.29).

$$\sum_{i=1}^{N_{SA}} \pi(s_i) = 1 \quad (4.29)$$

Using (4.26), the exact equilibrium distribution that fulfills (4.25) and (4.29), can be obtained as the solution to (4.30)

$$\begin{bmatrix} & & & & \\ & Q^T & & & \\ & & & & \\ 1 & \cdots & 1 & & \end{bmatrix} \times \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_{N_{SA}} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (4.30)$$

The dimension of Q is equal to the number of states N_{SA} in the Markov chain for a

certain SA policy. However, during the state-space generation one can also generate an auxiliary matrix $P = \{p_{i,j}\}$ for $i, j = 1 \dots N_{SA}$, whereas $p_{i,j} = 1$ (true) indicates that the transition from state i to state j is possible, and is set to zero (false) otherwise, according to the formation law. Because P is sparse, the elements of the matrix P can be stored in a memory in efficient way. Therefore P can be used to construct Q as a sparse matrix by only considering the $q_{i,j}$, if $p_{i,j} \neq 0$.

We note that the equation system in (4.30) is over-determined and any one of the first N_{SA} rows could be disregarded thus, turning the matrix on the LHS square, without changing the solution vector of the steady-state probabilities π_i . At first, we tried a method based on successive over-relaxation [86] to efficiently solve for the steady-state probabilities. However, we found that the conjugate gradient style solver method LSQR [85] is better suited for the sparse matrix problem (4.30) in our cases, as it apparently has better numerical properties, especially if Q is ill-conditioned. Furthermore, any direct and classical way of a closed form solution to solve (4.30), e. g. by calculating the inverse of Q , would not be an option, as an inversion or decomposition of a sparse matrix generally does not preserve the sparsity in intermediate calculations and thus increases complexity unnecessary.

4.3 Calculation of Blocking Probabilities

As we mentioned before, blocking can occur not only because of spectrum fragmentation, but also due to lack of available resources to serve an incoming connection request. In the following we will introduce the difference between *resource-* and *fragmentation-blocking*.

Blocking due to lack of resources is defined here as blocking that does not occur for reasons of fragmentation, but because the system does not have sufficient slots available to serve an incoming connection request. *Blocking due to fragmentation*, on the other hand, occurs when the system would have enough unused slots, but the free spectrum is scattered and, consequently, sufficient number of contiguous free slots to serve an incoming connection cannot be found.

For a given state $s_i \in S$, $P_c^R(s_i)$ denotes the probability that a specific request of type c , i.e., requiring n_c slots, cannot be satisfied due to lack of resource. It can be calculated based on the number of all free or idle slots (*#idleslots*) in a state s_i as

$$P_c^R(s_i) = \begin{cases} 1; & \text{if } n_c + 2n_G > \#idleslots \\ 0; & \text{otherwise} \end{cases} \quad (4.31)$$

By summing up these probabilities weighted by the probabilities of being in these states, we obtain the blocking probability $BP_c^R(s_i)$, *due to lack of resources*, for a specific

request type c , as

$$BP_c^R = \sum_{i=1}^N P_c^R(s_i) \cdot \pi(s_i) \quad (4.32)$$

The probability that a request of type c cannot be served due to spectrum fragmentation is denoted by $P_c^F(s_i)$. It can be calculated based on the size of the largest free block (e.g., size m slots) in the state s_i .

$$P_c^F(s_i) = \begin{cases} 1 - P_c^R(s_i); & \text{if } n_c + 2n_G > m \\ 0; & \text{otherwise} \end{cases} \quad (4.33)$$

Note that the definition (4.33) of $P_c^F(s_i)$ ensures that if there is blocking *due to lack of resources*, no contribution can be to $P_c^F(s_i)$. That is, if there is *resource-blocking*, then *fragmentation-blocking* is disregarded in any case, thus defining *blocking due to lack of resources* and *blocking due to spectrum fragmentation* as mutually exclusive events.

The *BP due to spectrum fragmentation*, for a specific incoming connection type c , can be calculated by summing up the probabilities $P_c^F(s_i)$ weighted by the steady-state probabilities of being in these states, as expressed in (4.34).

$$BP_c^F = \sum_{i=1}^N P_c^F(s_i) \cdot \pi(s_i) \quad (4.34)$$

If we assume the probability of occurrence per connection type being uniform, the total probability for *resource-* (4.35) and *fragmentation-blocking* (4.36) can be calculated by summing up BP_c^R and BP_c^F , respectively,, normalized by the number of request types.

$$BP^R = \frac{1}{C} \sum_{c=1}^C BP_c^R \quad (4.35)$$

$$BP^F = \frac{1}{C} \sum_{c=1}^C BP_c^F \quad (4.36)$$

Since *blocking due to lack of resources* and *due to fragmentation* are defined mutually exclusive, the overall blocking probability per request type c can be expressed by summing (4.32) and (4.34)

$$BP_c^{EOL} = BP_c^R + BP_c^F \quad (4.37)$$

and the overall blocking probability in an elastic optical link can be expressed as

$$BP^{EOL} = BP^R + BP^F \quad (4.38)$$

5 STATISTICAL ANALYSIS OF NODE- AND NETWORK-WISE OPERATION SCENARIO IN ELASTIC OPTICAL NETWORKS

This Chapter presents the results obtained with the proposed Markov-based framework. It begins by evaluating the statistical results for the node-wise modeling, in terms of spectrum fragmentation and BP. It then introduces the Monte Carlo simulation proposed for the analysis of large systems. The analytical results for a small scale elastic optical link are backed up by the simulated ones to show the validity of the model. Our aim, in this case, is to demonstrate the exactness of the analytical model, and also show a proof of concept for the proposed simulator to performance testing of large scale systems. Finally, the new definitions of blocking events are applied in the analysis of the dynamic resource provisioning in a network-wise operation of EONs.

5.1 Performance Results of Node-Wise Modeling

In this Section, we present the analytical results for the different SA schemes investigated in this thesis, in terms of BP and spectrum fragmentation, considering a node-wise modeling with $N = 30$ slots. Spectrum-slot requests are generated according to a Poisson process with arrival rates λ_c requests per time unit, and their duration is exponentially distributed with mean $1/\mu = 1$ time unit for all types of connection requests. The effective total average link load can be expressed in Erlang according to (5.1).

$$L = \sum_{c=1}^C \rho_c \cdot n_c \quad (5.1)$$

where ρ_c is obtained by multiplying λ_c by the mean holding time, which we assume equal to 1.

In the investigated scenario, we assume that there are three types of connection requests, which resource requirements correspond to $n_1 = 4$, $n_2 = 6$ and $n_3 = 8$ slots. Considering a slot bandwidth of 12.5 GHz and that the resource requirements include two guardband slots, see also Fig. 10, the assumed slot configurations correspond to exploitable transmission spectra of 25, 50 and 75 GHz. These bandwidth requests can be seen aligned with the variable line rate scenarios of 10, 40 and 100 Gb/s on elastic optical links, as investigated in [20].

We obtain the steady-state probabilities for different loads and SA policies and calculate BP and spectrum fragmentation in order to evaluate the relation between these two

performance measures.

As the Markov chain contains all the possible “spectrum configurations” that are represented by the states, it is possible to exactly identify the blocking that occurs due to the lack of resources and to differentiate it from the BP that occurs because of spectrum fragmentation. This is different from other analytical models in the literature [8] [27] that are not able to fully obtain all the possible spectrum configurations, and therefore, cannot achieve this level of differentiation. Thus, it is possible now to exactly show the gains, under the assumed system conditions, that can be achieved by defragmenting the elastic optical link. Such operation reduces blocking (probability) due to fragmentation to blocking only in case of lack of resources, which is often much lower as will be shown later in this section.

In Table 2, we present the complexity in terms of number of states for *Random-SA* policy, according to (4.21), for different number of slots and request sizes.

Table 2: Number of states for *Random-SA* policy

Number of Slots	Request Size (n_1, n_2, n_3)	Number of States
20		1319
30	(4, 6 and 8) slots	73150
40		4057374
20		5885
30	(3, 5 and 7) slots	652533
40		72353225

As can be seen, if *Random-SA* policy is to be analyzed, the complexity of the Markov model grows exponentially. However, the consideration and comparison of specific aspects of blocking, fragmentation and their interrelation and the findings for, required having analytical results for all (including *Random-*) SA policies under identical assumptions. Therefore, we had to limit the maximum number of slots in our investigation of an elastic optical link to $N = 30$. For larger N , numerical complexity exceeded our available computational resources. Due to the limitation that comes from a mismatch between the size of the simulation scenario (in terms of total number of slots) and the different sizes of connection requests, the levels of BP presented hereafter, are quite high and can be seen as not very realistic. Nevertheless we believe that they provide helpful insight in the behavior of BP and fragmentation for different network loads. In Section 5.2, we compare analytical to simulated results and find that our analytical results can describe the link in terms of blocking performance. Indeed, we realize the same relation between *blocking due to lack of resources* and *due to fragmentation* for the considered SA-policies over different link-loads for a simulated link with $N = 100$ slots.

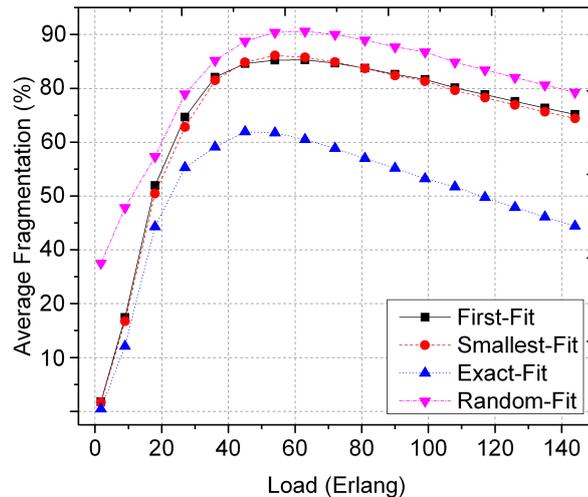


Figure 14: Average fragmentation based on (3.3).

5.1.1 Spectrum Fragmentation Analysis

Fig. 14 shows the average spectrum fragmentation according to (3.3), which relates the combination of both *internal* and *external*, as defined in (3.4) and (3.5), to the amount of currently free resources. For relatively low loads, the *EF* presents the lowest spectrum fragmentation compared with the other SA strategies described in Chapter 3. It should be noted that the fragmentation grows with increasing traffic load until it reaches the maximum and decreases with further increasing traffic load in an already highly loaded link. Consequently, fragmentation is not a unique quantity as function of load, i.e., a link can experience the same values of fragmentation for different levels of load. This behavior was also remarked in [8] and should be considered when used during planning or estimating the link performance in specific traffic conditions. Especially for low loads fragmentation results for *Random-Fit* policy is by far worse than for others SA policies. It does not even seem to approach a zero fragmentation measure in the considered scenarios.

The fragmentation metric defined in (3.2) relates the number of slots required for a certain connection and the number of such requests that could be accepted before blocking occurs, to the total number of free slots in the link. Using this metric, the fragmentation for different request sizes (in terms of numbers of contiguous slot required) is depicted in Fig. 15. Even though the average fragmentation performance in Fig. 14 indicates a clear ranking for the different SA policies, the fractional contributions of different request types show a non-homogeneous behavior in Fig. 15. The absolute contribution in case of *EF* for service requests requiring four slots is actually higher than for the other SA policies. However it shows significantly lower contributions to fragmentation for six and eight slot requests. The main contribution to fragmentation in case of *Random-Fit*, which shows the worst average fragmentation performance, is mainly due to requests for a large number of slots. No significantly higher contributions can be seen for small request sizes. It can be inferred that not only the spectrum allocation policy, but also the inter-

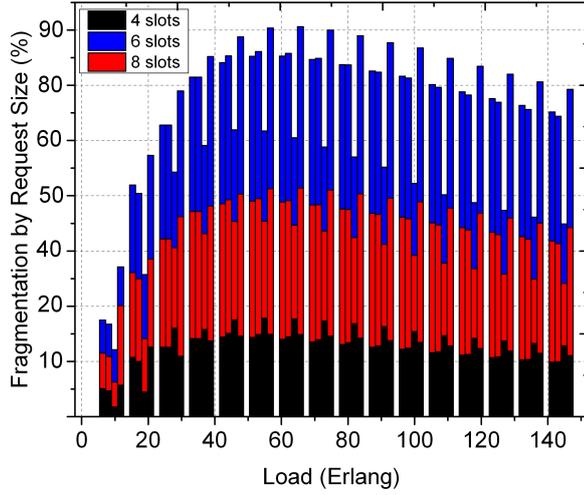


Figure 15: Spectrum fragmentation by request size according to (3.2) for FF , SF , EF and RND , left to right in each group.

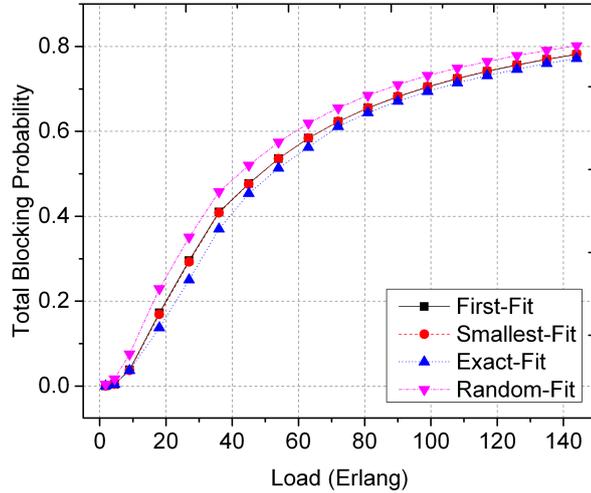


Figure 16: Total blocking probability for different SA policies based on (4.38).

play between different request sizes affect spectrum fragmentation and, consequently, the blocking performance.

5.1.2 Blocking Probability Analysis

Fig. 16 depicts the total blocking probability as a function of the load. As expected, the simulation results show that FF , SF and EF SA policies have better performance compared to RND . Under these conditions, the FF , SF and EF utilize spectrum more efficient due to allocation rules. We note that the total blocking probability of different SA strategies is very close as well as that the performance for higher load conditions seems to be inconsistent to the relative differences in SA policies' fragmentation performance. The reason for this will be explained in the next subsection.

At relatively low load, the blocking probability for FF , SF and EF increases moderately. At the higher load, the growth rate of BP increases significantly for all SA policies

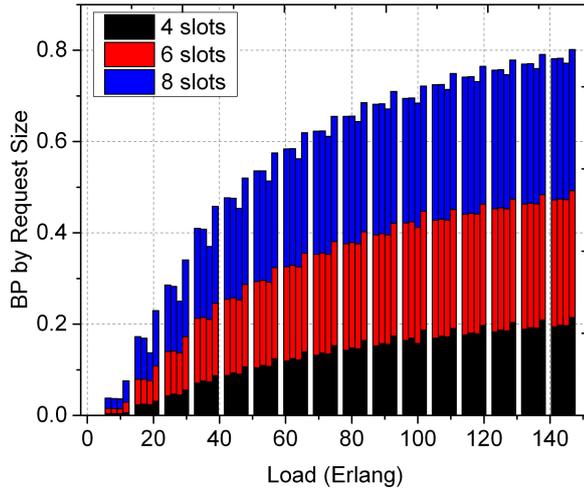


Figure 17: Blocking probability by request size according to (4.37) for FF , SF , EF and RND , left to right, in each group.

and is decreasing again under very high load conditions. The BP over load does not show any maximum for the considered loads. This is not what one would anticipate from the observation of fragmentation performance over load. With increased load another mechanism must affect blocking probability. Again this observation will be explained in the next subsection.

The impact on BP performance due to different spectral demands of connection requests considering different load conditions is shown in Fig. 17. Generally, the contribution to overall BP of requests increases with growing request size. For the considered scenarios, it appears that beyond a certain load (about 3.5 to 4 Erlang), the contribution to BP of larger requests becomes independent of the system load. BP for requests of four slots increases continuously with load, while BP for requests of eight slots shows saturation effects. It seems that for higher loads the blocking of larger requests becomes much more certain and therefore more or less constant, whereas for smaller requests the service availability is only reduced with increase of load. It seems there is no difference in behavior among the SA policies.

5.1.3 Resource- and Fragmentation-Blocking Analysis

The comparison between the impact of *fragmentation-blocking* and *resource-blocking* is shown in Fig. 18 and Fig. 19, respectively. Over the whole range of considered load, *Random-Fit* shows the worst performance for fragmentation blocking, while EF , SF and FF show similar performance with EF slightly better than the other two. When it comes to *resource-blocking*, *Random-Fit* exhibits lower BP compared to EF , SF and FF over a wide range of loads, with EF having highest *resource-blocking*. We will explain this behavior by comparing the best and worst performing SA strategies, i.e., EF and *Random-Fit*.

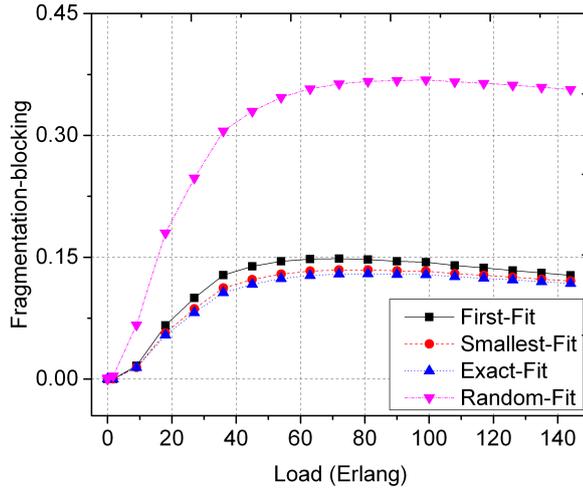


Figure 18: *Fragmentation-blocking* according to (4.36) for different SA policies and link loads.

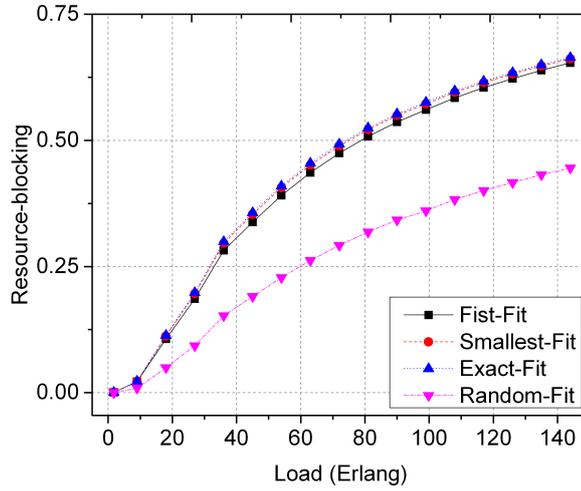


Figure 19: *Resource-blocking* based on (4.35) for different SA policies and link loads.

As can be seen in Fig. 14, *EF* has the lowest and *RND* the highest fragmentation for all considered loads. One can expect that if a SA policy is efficient, it should result in relatively low fragmentation. This expectation was confirmed by inspecting randomly chosen states generated by *EF* and *RND*. Because *EF* has the best fragmentation performance, this SA policy allows the most efficient spectrum utilization, leaving fewer slots unusable for serving further incoming requests. This on the other hand leads to more significant *BP due to lack of resources*. In contrast, *Random-Fit* is relatively inefficient in finding a suitable spectrum allocation. Therefore, *BP due to fragmentation* increases very fast with the load. This renders a certain amount of spectrum non-allocatable, making *BP due to fragmentation* dominating compared to *BP due lack of resources*, in particular for relatively low loads. This interrelation between *BP due to lack of resources* and *fragmentation* for *Random-Fit* is depicted in Fig. 20(a). It can be observed that beyond a certain link load *blocking due to lack of resources* eventually dominates. The *BP due to lack of resources* becomes also dominating for *FF* and *SF* policies, see Fig. 20(b) and

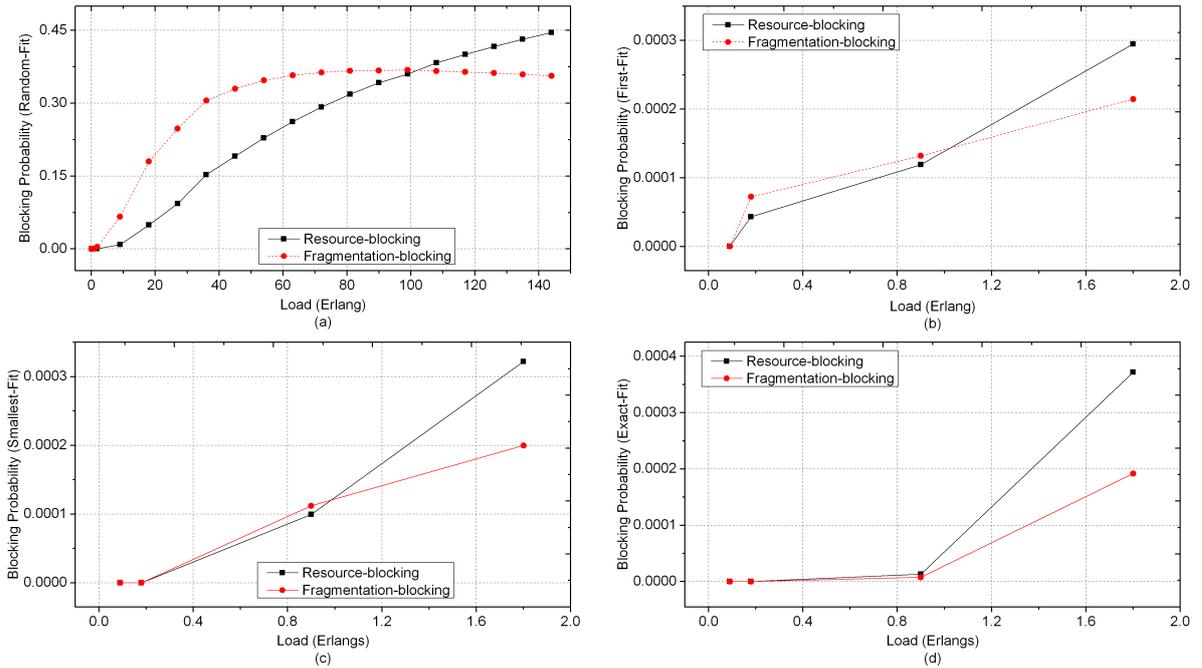


Figure 20: Contrast between (a) *RND*, (b) *FF*, (c) *SF* and (d) *EF* SA strategies when *resource-blocking* starts dominating over *fragmentation-blocking*.

Fig. 20(c), respectively, where the crossing point in BP for the two different reasons of blocking is already at relatively low load for and therefore not very expressed compared to *Random-SA* policy. In order to observe this, we had to reduce the considered range of load to 0.1 Erlang. For *EF*, Fig. 20(d) no crossing point could be found, which reflects the efficiency in spectrum allocation of this SA strategy. Comparing the magnitudes of *BP due to lack of resources* and *fragmentation*, one can see from Fig. 20(a) that early defragmentation for *Random-Fit* can indeed provide gains in terms of reduction of overall blocking probability. As can be seen from Fig. 20(b), Fig. 20(c) and Fig. 20(d) at load below 1 Erlang *BP due to lack of resources* is similar for all SA policies. Since for the other SA policies, namely *FF*, *SF* and *EF*, the *lack of resources* becomes almost immediately the dominating driver for blocking, defragmentation would provide gains only at very low loads and of limited magnitude.

5.2 Monte Carlo Simulation

In this Section, we present results based on simulations of an elastic optical link. We compare the analytical results for $N = 30$ slots with the simulated ones under identical assumption and it is shown that analytical and simulation results are very similar. Then, we present probabilities for *blocking due to lack of resources* and *blocking due to fragmentation* for a large scale elastic optical link with $N = 100$ slots obtained by simulations.

Results for different link-loads were derived using an event-driven simulator based on Poisson traffic with arrival rates λ_c requests per time unit, and exponentially distributed

Table 3: BP of different SA policies for small scale EOL (%)

SA-Policy	Load (Erlang)							
	18		36		72		108	
	Exact	Sim.	Exact	Sim.	Exact	Sim.	Exact	Sim.
FF	16.57	16.56	39.69	39.69	62.29	62.28	72.42	72.42
SF	16.21	16.21	39.48	39.41	62.34	62.34	72.52	72.52
EF	16.05	16.05	39.23	39.23	62.17	62.17	72.41	72.41
RND	22.93	22.04	45.78	44.52	65.53	65.51	74.89	74.85

Table 4: *Resource- and fragmentation-blocking* of different SA policies for small scale EOL (%)

SA-Policy	Load (Erlang)															
	18				36				72				108			
	Ex.		Sim.		Ex.		Sim.		Ex.		Sim.		Ex.		Sim.	
	Res.	Frag.	Res.	Frag.	Res.	Frag.	Res.	Frag.	Res.	Frag.	Res.	Frag.	Res.	Frag.	Res.	Frag.
FF	10.19	6.38	10.19	6.38	27.13	12.56	27.13	12.56	47.47	14.82	47.46	14.82	58.44	13.98	68.44	13.98
SF	10.72	5.49	10.72	5.49	28.50	10.97	28.5	10.95	48.91	13.43	48.91	13.43	59.53	12.99	59.53	12.99
EF	10.83	5.22	10.83	5.22	28.81	10.42	28.82	10.42	49.24	12.92	49.25	12.92	59.80	12.61	59.81	12.61
RND	4.94	17.99	4.86	17.18	15.25	30.53	14.76	29.76	29.19	36.34	29.45	36.06	38.29	36.60	38.64	36.21

holding times with mean holding time equal to one time unit. We generally assume heterogeneous distribution of connection request types. In case of arrival events, the slot assignment to connection requests is performed according to the considered SA-policy. If, depending on the status of slot occupancy, a connection request can be served, it is allocated and the link state is updated accordingly. In case, a connection request is rejected, the link state remains unchanged and it is investigated, whether it is for reasons of *blocking due to lack of resources* or *due to spectrum fragmentation*. The appearance of each blocking event type is tracked by individual counters. In case of termination of a formerly allocated connection, the occupied slots are released, resulting in a corresponding link state change. The blocking probabilities are derived as the average blocking event appearance rates relative to the number of arrival events.

Each data point is based on averaging results from twenty independent simulations of ten million arrival events each. Table 3 and Table 4 show the comparison of blocking probabilities obtained by simulations and by using our analytical model for small scale EOL. It can be seen that the results of the analytical model and the simulations are very close. Simulation results for blocking probability considering $N = 100$ slots are provided in Fig. 21, Fig. 22 as well as in Fig. 23.

For this simulation, three sizes of connection requests are assumed, i.e., $n_1 = 5$, $n_2 = 10$ and $n_3 = 15$ slots. The total blocking probabilities for *FF*, *SF* and *EF* SA-policy appear

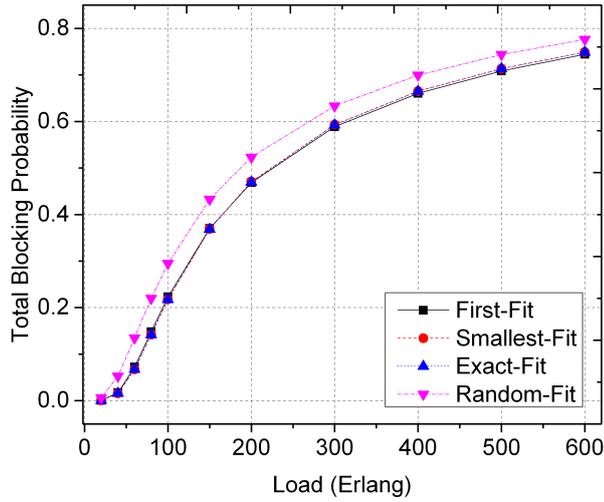


Figure 21: Total blocking probability for different SA policies based on (4.38) for large scale EOL.

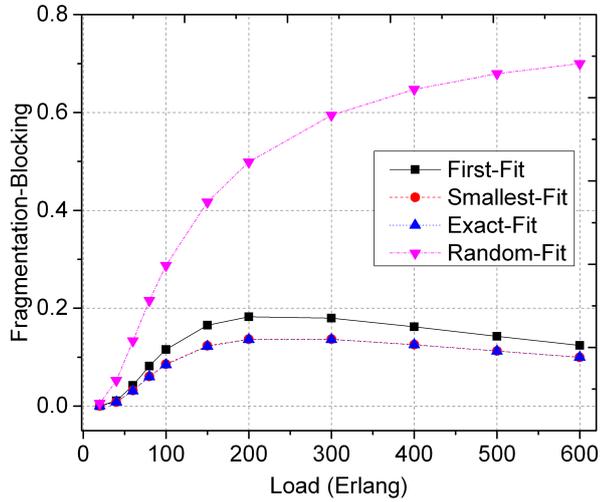


Figure 22: *Fragmentation-blocking* based on (4.36) for different SA policies and link loads for large scale EOL.

to be identical. *Random-SA* policy performs clearly inferior. However, observing the different contribution from *fragmentation-* and *resource-blocking* in Fig. 22 and Fig. 23 respectively, we see that blocking for *Random-SA* policy is overwhelmingly dominated by blocking due to fragmentation; *EF* SA-policy shows the lowest level of *fragmentation-blocking*, which gives indication of a better efficiency in terms of resource allocation. Overall, the analytical result for the small scale and simulated results for the larger scale link scenario show very much the same trends of blocking performance. This suggests that analysis of the dynamic behavior of a small scale EOL can already provide useful information about spectrum occupancy and performance measures in terms of blocking and fragmentation of larger scale EOL.

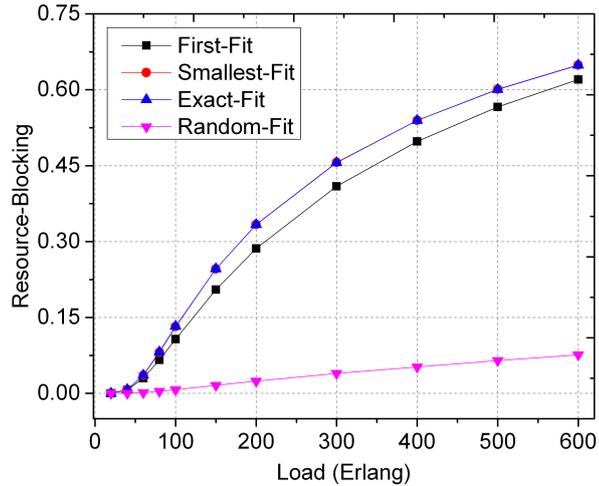


Figure 23: *Resource-blocking* according to (4.35) for different SA policies and link loads for large scale EOL.

5.3 Performance Results of Network-Wise Modeling

In this Section, the results for network-wise BP for total BP and the different blocking events, i.e., *resource-* and *fragmentation-blocking*, of different operation modes are evaluated over a simple network topology, for large scale parameters.

5.3.1 Description of the Network

The network resembles a variant of broadcast-and-select networking. Its architecture is depicted in Fig. 24. A practical implementation of such network architecture using *spectrum selective switch* (SSS) may be considered for enabling the elastic optical nodes. It can be implemented using splitters at the input ports that generate copies of the incoming signals, which are subsequently filtered by the SSS in order to select the required signals at the output. In addition to that, the add/drop network may implement colorless, directionless and contentionless elastic add/drop functionality in order to enable full flexibility in the network. However, an in-depth analysis of the network constraints as well as its main characteristics is not the scope of this thesis.

The relation of the end-to-end nodes of a certain connection request and the nodes over which this connection is supposed to be established are given in Table 5. Not all end-to-end connection are tabulated, since a connection from start-node i to end-node j is attributed the same path as a connection from node j to i .

5.3.2 Description of the Simulation

The network-wise EON scenarios are evaluated based on an event-driven simulation, which corresponds to an extension of the same simulator as described in Section 5.2. The EON

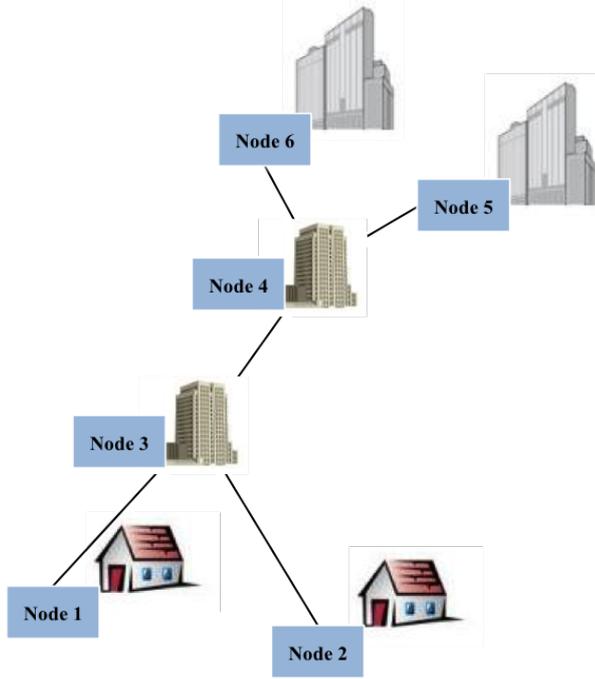


Figure 24: Illustration of a simple *broadcast-and-select* network topology.

Table 5: Relation of end-to-end link definition and involved nodes for the network in Fig. 24

End-to-End Link	Involved Nodes
1 – 2	1-3-2
1 – 3	1-3
1 – 4	1-3-4
1 – 5	1-3-4-5
1 – 6	1-3-4-6
2 – 3	2-3
2 – 4	2-3-4
2 – 5	2-3-4-5
2 – 6	2-3-4-6
3 – 4	3-4
3 – 5	3-4-5
3 – 6	3-4-6
4 – 5	4-5
4 – 6	4-6
5 – 6	5-4-6

simulator generates Poisson traffic with mean holding time equal to one time unit and serves network lightpath requests one-by-one.

For the performed simulations, the number of required spectrum slots of each lightpath is randomly selected from the set of $(n_1, n_2, n_3) = (5, 10, 15)$ slots. The number of initially available slots on each node is assumed to be $N = 300$ [27]. The simulation for each

load comprises 10^6 arrival events; each simulation is repeated five times and results are averaged.

The procedure to derive the BP for our network evaluation is equivalent to the one in the link evaluation, Section 5.2. In case, a connection request is rejected, the network state remains unchanged and it is investigated, whether the blocking of the connection request occurs for reasons of *blocking due to lack of resources* or *due to spectrum fragmentation*, considering all nodes involved. The appearance of each blocking event type is tracked by individual counters and related to the number of arrival events.

5.3.3 Results

Figure 25 shows the total blocking probability for $N = 300$ as a function of the network load. As can be observed, essentially no blocking occurs in this network case below a load of 300, for any considered SA policy. *First-Fit*, *Smallest-Fit* and *Exact-Fit*, on the other hand, show very low blocking even below load 600, which is a significantly higher load for such observation, if we compare with their blocking performance in the analysis of the large EOL (see Fig. 21), where the overall blocking becomes visible at load 50, i.e., when the system load is half the overall resources (slots) available. Such aspect should be further investigated taking into account that, as studied in this thesis, blocking probability might have different contributions. In addition to that, in case of the network evaluation, not every node is involved in every end-to-end connection request. For instance, as one can see from Table 5, in a connection from node 1 to node 5, only nodes 1, 3, 4, 5 are required. Therefore, different demands of connection request do not affect all network nodes evenly.

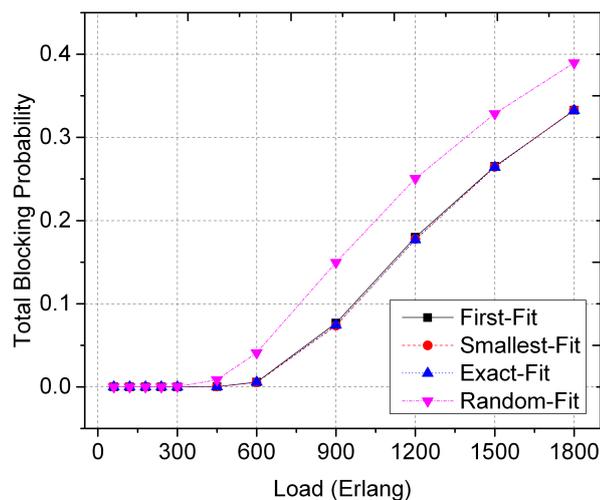


Figure 25: Total blocking probability for an elastic network-wise scenario as a function of the load, assuming $N = 300$ slots.

On the other side, in order to allocate an incoming connection, a suitable slot configuration must be provided by all involved nodes simultaneously. i.e., both *spectrum contiguity* and

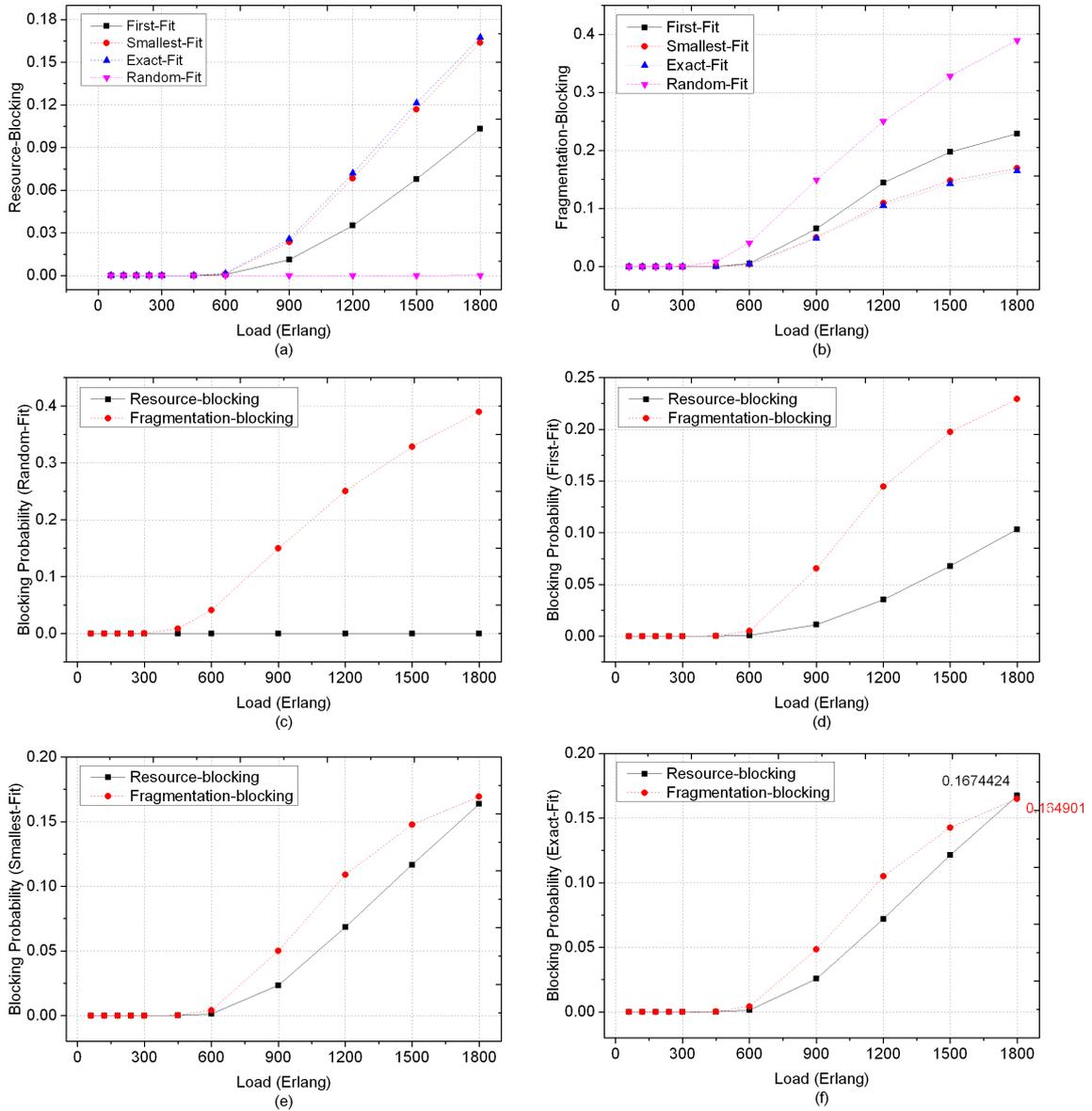


Figure 26: (a) *Fragmentation-blocking* and (b) *resource-blocking* for an elastic network-wise scenario, assuming $N = 300$ slots; contrast between (c) *RND*, (d) *FF*, (e) *SF* and (f) *EF* SA strategies when *resource-blocking* starts dominating *fragmentation-blocking*.

spectrum continuity constraints must be fulfilled to allocate an optical path, making it mode difficult to find one.

As can be seen in Fig. 25, *Random-SA* policy performs significantly worse than *FF*, *SF* and *EF*, and by analyzing Fig. 26(a) and Fig. 26(b), respectively, it is clear that blocking for *RND* is solely contributed by *fragmentation-blocking* (see also Fig. 26(c)). Although, in Fig. 25 we can observe the same blocking performance for *FF*, *SF* and *EF*, in Fig. 26(b) we can clearly notice that *fragmentation-blocking* in case of *FF* is higher compared to *SF* and *EF*, and the opposite for *resource-blocking* (see Fig. 26(a)). It shows that, for the considered loads, the effect of fragmentation is much more pronounced in case of *First-Fit* SA policy, opposite to *SF* and *EF*. Such aspect is contrary to what can be observed by

comparing corresponding results for *fragmentation-* and *resource-blocking* for small (see Fig. 18 and Fig. 19) as well as large scale (see Fig. 22 and Fig. 23) link-modeling. It seems that the strategy of allocating slots for connection requests always in the first available block leads to higher fragmentation from the network-wise point of view, compared to the fragmentation observed for individual network sections.

Figure 26(e) and Fig. 26(f) depict the relation between *fragmentation-* and *resource-blocking* for *Smallest-Fit* and *Exact-Fit*, respectively. As we can see, both SA policies show the same behavior as found in our evaluation for the small scale link-modeling (see Fig. 20(c) and (d), respectively); at first blocking is dominated due to spectrum fragmentation, and eventually dominated due to *lack of available resources* in the network. Note that the load for the network scenario needs to be significantly higher than for the link-model with 30 slots. On the other hand, the considered request sizes in case of the network-modeling are much more granular, relative to the amount of resources, when compared to the link-modeling.

So far, *Smallest-Fit* and *Exact-Fit* have always shown relatively similar performance for higher loads, whether in case of link- or network-modeling. This is an unexpected performance for the reason that, if there is no exact available block of resources, *EF* follows the complete opposite strategy than *SF*, choosing the largest available block to allocate a connection request. Therefore, further investigations should be carried out in order to evaluate the underlying reasons for such outcome.

6 CONCLUSIONS AND FUTURE WORKS

6.1 Conclusion

This thesis addressed the resource provisioning paradigm in Elastic Optical Networks (EONs) with dynamic traffic conditions. The work focused on the spectrum fragmentation problem, and its effect on requests' blocking probability. We also considered the underlying reasons for different types of blocking events and draw conclusions on effective defragmentation gains and efficiency of dynamic spectrum allocation schemes, providing a step in network-wise evaluation.

We presented a comprehensive review of the evolution of optical nodes and networks showing that flexibility has increased with each evolutionary stage so that additional capacity and functionality are supported. We also revised the state-of-the-art of the experimental techniques suitable for the generation of multi-Tb/s channels and explain how the super-channel principle can support elastic optical networking. High-capacity transmission experiments were also surveyed to determine optical node scalability requirements and whether current optical design approaches are suitable to implement high-capacity networks. We also explain reasons for the need for elastic resource allocation and adaptive optical infrastructure with efficient multi-granular support.

While addressing the elastic resource allocation paradigm in EONs, we proposed a novel SA policy, so called *Exact-Fit* and compared it with existing SA strategies in order to analyze the efficiency of different dynamic spectrum allocation schemes. We showed that smart spectrum allocation policies that consider size of the allocated block relative to the size of the available spectrum block, perform better compared to random or sequential approaches. We have also shown that, if it is not possible to find an exactly fitting available spectrum block, it is suitable to accommodate the largest fitting block in order to not leave behind too small blocks that would lead to fragmentation. It is worth mentioning that the presented analytical way to study the stochastic process corresponding to the dynamic spectrum occupancy in elastic optical links, was the first presented attempt to calculate blocking probability together with fragmentation ratio.

In our study on the effect of fragmentation on requests' blocking probability, we introduced new definitions for blocking that differentiate the reasons for blocking events. In this respect, we derived a Markov-based model that proposes a new way of assessing the SA schemes, and a new way to analyze the dependency between the blocking probability and fragmentation. In this respect, a new accommodated fragmentation metric was introduced, allowing differentiating minute (i.e. small) variations of spectrum occupancy. This new approach made it possible to the investigation of how different request sizes can contribute to blocking-probability and fragmentation of spectral resources. We

found that the interplay between different request sizes affects bandwidth fragmentation and, consequently, the blocking performance. In addition to that, our probabilistic model confirms that when the spectrum becomes more fragmented, the blocking probability is increased and the advantage of elastic resource allocation is reduced.

Since our analytical model can exactly calculate the values for *resource-* and *fragmentation-blocking*, the trend of the impact of spectrum fragmentation due to increased load on BP can be analyzed. It offers, therefore, a new way to forecast how much BP will change with spectral fragmentation for a given SA policy. This analysis can demonstrate at which point the defragmentation would be able to postpone the network upgrade with more resources, which can be hopefully developed further to give useful and directly helpful message to network operators.

Our study on the effect of fragmentation on requests' blocking probability also concludes that only the analysis of fragmentation ratio is not sufficient to make the decision when it may be beneficial to perform the defragmentation process, since the fragmentation ratio not always reflects the blocking performance. This is because the existing definitions of fragmentation ratio are not directly correlated with the network blocking performance. Therefore, the decision for starting a defragmentation process to reduce the blocking probability should be based on the relation between blocking probability, spectrum fragmentation and requested bandwidths.

To address the problem of statistically analyze large scale systems, we developed a Monte Carlo simulation in order to provide expected blocking probability for large scale elastic node- and network-wise operation scenarios. The results of the analytical model were backed up with the simulations and were very close to each other. Simulation results for a large scale elastic links scenario show exactly the same trends of blocking performance as the analytical results, implying that analysis of the dynamic behavior of a small scale EOL can already provide useful information about spectrum occupancy and performance measures in terms of blocking and fragmentation of larger scale EOL.

Finally, we believe that our Markov-based framework can also serve as a tool for evaluation of different definitions for fragmentation ratio in terms of their usefulness as the network performance metric. Our Monte Carlo simulation, on the other hand, can be easily adapted to analyze more complex network topologies. From the experiences on running different instances, it can be said that our both implementations are quite efficient, showing a good computational run time when compared to existing related works in the literature.

6.2 Future Research

In terms of future work for dynamic resource allocations, the Markov-based framework proposed in this thesis can be extended for serving time-varying traffic fluctuations in

an EON that enables the dynamic spectrum sharing among connections. Such analysis is relevant because demands are not only dynamic in terms of arrival and duration but they are also dynamic in terms of bandwidth. For instance, a service provider may request more bandwidth from the network in some time periods to support the data backup service during night hours or to support video-on-demand services during evening hours. In conventional WDM networks, whenever significant traffic changes are observed or expected, additional wavelength are provisioned. On contrary, in elastic optical networks, the allocated optical spectrum of a lightpath is tailored to the actual width of the transmitted signal. In such a case, an off-line RSA is not, of course, adequate to address this dynamic. That is the reason that, recently, research has advanced from off-line to on-line RSA. Therefore, the statistical analysis of more complex approaches considering the elastic time/spectrum allocation, i.e., the expansion (or contraction) of the spectrum when the required bit rate of an demand increases (or decreases), should be carried out.

Another area for improvements lies in the statistical analysis of optical grooming techniques, which have been recently proposed to enabling the aggregation and distribution of the traffic directly at the optical layer. Thus, multiple low capacity demands can be grouped together in the same transceiver. Approaches that do not require the use of guardband or multiple light sources in flexible transceivers could be investigated in order to analyze the significant savings achieved in terms of transceivers and spectrum usage. The support of such functionality imposes further requirements on the architecture of the optical switches. For example, it leads to adoption of *broadcast-and-select* structures, strengthening even more our analysis in network architectures where no optical switching is involved.

Further extensions of our work may include the integration with parallel and distributed approaches, in order to reduce computational complexity of solving large models. A distributed implementation of an iterative method will enable very large models to be solved on clusters of modest sizes by reducing the main memory (RAM) requirements. An increasing trend in the area of distributed and grid computing has been observed. Major developments have been made in the area of communication technologies, both in software and hardware. Disk technology has also benefited from the introduction of RAIDs. Therefore, we anticipate that a combination of parallel and iterative techniques will play an important role in the analysis of large CTMCs.

Possible future research topics in the area of design and network optimization in elastic optical networking may include the investigation of networks that provides a biggest path options due to add node interconnections for reasons of resilience. As services impose high availability requirements, it is of vital importance to design resilient networks. Resilience can be provided via dedicated and shared protection schemes. In contrast to dedicate protection schemes, shared protection schemes allow sharing of the protection resources among connection with disjoint working resources, resulting in superior spectrum effici-

ency and best performance in terms of energy efficiency. Another interesting aspect to address in design and network optimization would be the evaluation in how nodes should be (ideally) interconnected, i.e., what kind of link should be provided between the nodes in order to reduce the overall blocking in the network.

Bibliography

- [1] Cisco visual networking index: Forecast and methodology, 2013–2018, June 2014.
- [2] W. Shieh and I. Djordjevic. *OFDM for Optical Communications*. Academic Press, Elsevier, 1 edition, October 2009.
- [3] R. S. Tucker. Optical packet switching: A reality check. *Optical Switching and Networking*, 5(1):2–9, June 2008.
- [4] M. Jinno, H. Takara, B. Kozicki, Yukio Tsukishima, Y. Sone, and S. Matsuoka. Spectrum-efficient and scalable optical path network: Architecture, benefits, and enable technologies. *Communications Magazine, IEEE*, 47(11):66–73, November 2009.
- [5] S. L. Jansen, I. Morita, K. Forozesh, S. Randel, Dirk van den Borne, and H. Tanaka. Optical ofdm, a hype or is it for real? In *Communication, 2008. ECOC 2008. 34th European Conference on*, pages 1–4, September 2008.
- [6] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri. On the performance of nyquist-wdm terabit superchannels based on pm-bpsk, pm-qpsk, pm-8qam or pm-16qam subcarriers. *Lightwave Technology, Journal of*, 29(1):53–61, January 2011.
- [7] D. J. Geisler, N. K. Fontaine, R. P. Scott, T. He, L. Paraschis, O. Gerstel, J. P. Heritage, and S. J. Yoo. Bandwidth scalable, coherent transmitter based on the parallel synthesis of multiple spectral slices using optical arbitrary waveform generation. *Optics Express*, 19(9):8242–8253, April 2011.
- [8] Y. Yu, J. Zhang, Y. Zhao, H. Li, Y. Ji, and W. Gu. Exact performance analytical model for spectrum allocation in flexible grid optical networks. *Optical Fiber Technology*, 20(2):75–83, January 2014.
- [9] ITU-T. Transmission systems and media, digital systems and networks, February 2012.
- [10] M. Presi, F. Bottoni, R. Corsini, G. Cossu, and E. Ciaramella. All dfb-based coherent udwdm pon with 6.25 ghz spacing and a power budget power budget. *Photonics Technology Letters, IEEE*, 26(2):107–110, January 2014.
- [11] I. Kaminow, T. Li, and A. Willinger. *Optical Fiber Telecommunications*, volume VIB. Academic Press, Elsevier, 6 edition, May 2013.
- [12] O. Gerstel, M. Jinno, A. Lord, and S. J. B. Yoo. Elastic optical networking: a new dawn for the optical layer? *Communications Magazine, IEEE*, 50(2):12–20, February 2012.

- [13] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano. Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network. *Communications Magazine, IEEE*, 48(8):138–145, August 2010.
- [14] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos. Elastic bandwidth allocation in flexible ofdm-based optical networks. *Lightwave Technology, Journal of*, 29(9):1354–1366, May 2011.
- [15] M. Klinkowski and K. Walkowiak. Routing and spectrum assignment in spectrum sliced elastic optical path network. *Communications Letters, IEEE*, 15(8):884–886, August 2011.
- [16] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos. Dynamic bandwidth allocation in flexible ofdm-based networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.
- [17] X. Wan, N. Hua, and X. Zheng. Dynamic routing and spectrum assignment in spectrum-flexible transparent optical networks. *Optical Communications and Networking, IEEE/OSA Journal of*, 4(8):603–613, August 2012.
- [18] Y. Wang, J. Zhang, Y. Zhao, J. Wang, and W. Gu. Routing and spectrum assignment by means of ant colony optimization in flexible bandwidth networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [19] A. Castro, L. Velasco, M. Ruiz, M. Klinkowski, J. P. Fernández-Palacios, and D. Careglio. Dynamic routing and spectrum (re)allocation in future flexgrid optical networks. *Computer Networks*, 56(12):2869–2883, August 2012.
- [20] I. Stiakogiannakis, E. Palkopoulou, D. Klonidis, O. Gerstel, and I. Tomkos. Dynamic cooperative spectrum sharing and defragmentation for elastic optical networks. *Optical Communications and Networking, IEEE/OSA Journal of*, 6(3):259–269, March 2014.
- [21] R. Wang and B. Mukherjee. Provisioning in elastic optical networks with non-disruptive defragmentation. *Lightwave Technology, Journal of*, 31(15):2491–2500, August 2013.
- [22] D. J. Geisler, R. Proietti, Y. Yin, R. P. Scott, N. K. Cai, X. Fontaine, L. Paraschis, O. Gerstel, and S. J. B. Yoo. Experimental demonstration of flexible bandwidth networking with real-time impairment awareness. *Optics Express*, 19(26):B736–B745, December 2011.

- [23] X. Cai, K. Wen, R. Proietti, Y. Yin, D. J. Geisler, R. P. Scott, C. Qin, L. Paraschis, O. Gerstel, and S. J. B. Yoo. Experimental demonstration of adaptive combinational qot degradation restoration in elastic optical networks. *Lightwave Technology, Journal of*, 31(4):664–671, February 2013.
- [24] K. Wen, X. Cai, Y. Yin, D. J. Geisler, R. Proietti, R. P. Scott, N. K. Fontaine, and S. J. B. Yoo. Adaptive spectrum control and management in elastic optical networks. *Selected Areas in Communications, IEEE Journal on*, 31(1):39–48, January 2013.
- [25] S. Yang and F. Kuipers. Impairment-aware routing in translucent spectrum-sliced elastic optical path networks. In *Networks and Optical Communications (NOC), 2012 17th European Conference on*, pages 1–6, June 2012.
- [26] H. Beyranvand and J. A. Salehi. A quality-of-transmission aware dynamic routing and spectrum assignment scheme for future elastic optical networks. *Lightwave Technology, Journal of*, 31(18):3043–3054, September 2013.
- [27] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos. Time-varying spectrum allocation policies and blocking analysis in flexible optical networks. *Selected Areas in Communications, IEEE Journal on*, 31(1):13–25, January 2013.
- [28] L. Peng, C. H. Youn, and C. Qiao. Theoretical analyses of lightpath blocking performance in co-ofdm optical networks with/without spectrum conversion. *Communications Letters, IEEE*, 17(4):789–792, April 2013.
- [29] H. Beyranvand, M. Maier, and J. A. Salehi. An analytical framework for the performance evaluation of node- and network-wise operation scenarios in elastic optical networks. *Communications, IEEE Transactions on*, 62(5):1621–1633, May 2014.
- [30] W. Shi, Z. Zhu, M. Zhang, and N. Ansari. On the effect of bandwidth fragmentation on blocking probability in elastic optical networks. *Communications, IEEE Transactions on*, 61(7):2970–2978, July 2013.
- [31] T. Takagi, H. Hasegawa, K. Sato, Y. Sone, A. Hirano, and M. Jinno. Disruption minimized spectrum defragmentation in elastic optical path networks that adopt distance adaptive modulation. In *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, September 2011.
- [32] X. Wang, Q. Zhang, I. Kim, P. Palacharla, and M. Sekiya. Utilization entropy for assessing resource fragmentation in optical networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.

- [33] X. Yu, J. Zhang, Y. Zhao, L. Peng, Y. Bai, D. Wang, and X. Lin. Spectrum compactness based defragmentation in flexible bandwidth optical networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [34] R. Wang and B. Mukherjee. Spectrum management in heterogeneous bandwidth optical networks. *Optical Switching and Networking*, 11:83–91, October 2014.
- [35] A. Rosa, C. Cavdar, S. Carvalho, J. Costa, and L. Wosinska. Spectrum allocation policy modeling for elastic optical networks. In *Proc. 9th International Conference on High Capacity Optical Networks and Enabling Technologies (HONET)*, pages 242–246, 2012.
- [36] A. Rosa, Wiatr. P., C. Cavdar, S. V. Carvalho, J. C. W. A. Costa, and L. Wosinska. Statistical analysis of blocking probability and fragmentation based on markov modeling of elastic optical spectrum allocation on fiber link. *Optics Communication*, 2015.
- [37] C. Trembly, F. Gagnon, B. Chatelain, E. Bernier, and M. P. Belanger. Filterless optical networks: A unique and novel passive wan network solution. In *Proc. OECC/IOOC*, pages 1–2, 2009.
- [38] G. Keiser. *Optical Fiber Communication*. McGraw-Hill Science in Electrical and Computer Engineering, 4 edition, September 2010.
- [39] R. Ramaswami, K. N. Sivarajan, and G. Sasaki. *Optical Networks: a practical perspective*. Morgan Kaufmann, 3 edition, July 2009.
- [40] P. N. Ji and Y. Aono. Colorless and directionless multi-degree reconfigurable optical add/drop multiplexers. In *Wireless and Optical Communications Conference (WOCC), 2010 19th Annual*, pages 1–5, May 2010.
- [41] A. Sano, E. Yamada, H. Masuda, E. Yamazaki, T. Kobayashi, E. Yoshida, Y. Miyamoto, R. Kudo, K. Ishihara, and Yasushi Takatori. No-guard-interval coherent optical ofdm for 100-gb/s long-haul wdm transmission. *Lightwave Technology, Journal of*, 27(16):3705–3713, August 2009.
- [42] D. Qian, M. Huang, E. Ip, Y. Huang, Y. Shao, J. Hu, and T. Wang. 101.7-tb/s (370×294-gb/s) pdm-128qam-ofdm transmission over 3×55-km ssmf using pilot-based phase noise mitigation. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.

- [43] M. Salsi, H. Mardoyan, P. Tran, C. Koebele, E. Dutisseuil, G. Charlet, and S. Bigo. 155×100gbit/s coherent pdm-qpsk transmission over 7,200km. In *Optical Communication, 2009. ECOC '09. 35th European Conference on*, volume 2009-Supplement, pages 1–2, September 2009.
- [44] T. J. Xia, G. A. Wellbrock, Y. Huang, M. Huang, E. Ip, P. N. Ji, D. Qian, A. Tanaka, Y. Shao, T. Wang, Y. Aono, and T. Tajima. 21.7 tb/s field trial with 22 dp-8qam/qpsk optical superchannels over 1,503-km of installed ssmf. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [45] S. Randel, R. Ryf, A. H. Gnauck, M. A. Mestre, C. Schmidt, R. Essiambre, P. J. Winzer, R. Delbue, P. Pupalais, A. Sureka, Y. Sun, X. Jiang, and R. L. Lingle. Mode-multiplexed 6×20-gbd qpsk transmission over 1200-km dgd-compensated few-mode fiber. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [46] B. Zhu, T. Taunay, M. Fishteyn, X. Liu, S. Chandrasekhar, and M. F. Yan. Space-, wavelength-, polarization-division multiplexed transmission of 56-tb/s over a 76.8-km seven-core fiber. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.
- [47] J. et al. Sakaguchi. 19-core fiber transmission of 19×100×172-gb/s sdm-wdm-pdm-qpsk signals at 305tb/s. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [48] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, R. Essiambre, P. J. Winzer, D. W. Peckham, A. McCurdy, and R. Lingle. Space-division multiplexing over 10 km of three-mode fiber using coherent 6 × 6 mimo processing. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.
- [49] K. S. Abedin, T. F. Taunay, M. Fishteyn, M. F. Van, B. Zhu, J. M. Fini, E. M. Monberg, F. V. Dimarcello, and P. W. Wisk. Amplification and noise properties of an erbium-doped multicore fiber amplifier. *Optics Express*, 19(17):16715–16721, August 2011.
- [50] Y. Yung, S. Alam, Z. Li, A. Dhar, D. Giles, I. Giles, J. Sahu, L. Gruner-Nielsen, F. Poletti, and D. J. Richardson. First demonstration of multimode amplifier for spatial division multiplexed transmission systems. In *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, September 2011.

- [51] N. Bai, E. Ip, T. Wang, and G. Li. Multimode fiber amplifier with tunable modal gain using a reconfigurable multimode pump. *Optics Express*, 19(17):16601–16611, August 2011.
- [52] G. Gavioli, E. Torrenco, G. Bosco, A. Carena, V. Curri, V. Miot, P. Poggiolini, M. Belmonte, F. Forghieri, C. Muzio, S. Piciaccia, A. Brinciotti, A. La Porta, C. Lezzi, S. Savory, and S. Abrate. Investigation of the impact of ultra-narrow carrier spacing on the transmission of a 10-carrier 1tb/s superchannel. In *Optical Fiber Communication (OFC), collocated National Fiber Optic Engineers Conference, 2010 Conference on (OFC/NFOEC)*, pages 1–3, March 2010.
- [53] H. Nyquist. Certain topics in telegraph transmission theory. *American Institute of Electrical Engineers, Transactions of the*, 47(2):617–644, April 1928.
- [54] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IEEE*, 86(2):447–457, February 1998.
- [55] M. A. Soto, M. Alem, M. A. Shoaie, A. Vedadi, C. Brès, L. Thévenaz, and T. Schneider. Optical sinc-shaped nyquist pulses of exceptional quality. *Nature Communication*, 4, December 2013.
- [56] B. et al. Kozicki. Experimental demonstration of spectrum-sliced elastic optical path network (slice). *Optics Express*, 18(21):22105–22118, October 2010.
- [57] O. Rival and A. Morea. Cost-efficiency of mixed 10-40-100gb/s networks and elastic optical networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.
- [58] C. T. Politi, V. Anagnostopoulos, C. Matrakidis, A. Stavdas, A. Lord, V. Lopez, and J. P. Fernandez-Palacios. Dynamic operation of flexi-grid ofdm-based networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [59] A. Morea, A. F. Chong, and O. Rival. Impact of transparent network constraints on capacity gain of elastic channel spacing. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.
- [60] S. L. Jansen, G. D. Khoe, H. De Waardt, S. Spalter, H. E. Escobar, M. Sher, Dirk Woll, and D. Zhou. 10 gbit/s, 25 ghz spaced transmission over 800 km using dispersion compensation modules. In *Optical Fiber Communication Conference, 2004. OFC 2004*, volume 2, page 3, February 2004.

- [61] H. Suzuki, M. Fujiwara, N. Takachio, K. Iwatsuki, T. Kitoh, and T. Shibata. 12.5 ghz spaced 1.28 tb/s (512-channel x 2.5 gb/s) super-dense wdm transmission over 320 km smf using multiwavelength generation technique. *Photonics Technology Letters, IEEE*, 14(3):405–407, March 2002.
- [62] M. Jinno, K. Yonenaga, H. Takara, K. Shibahara, S. Yamanaka, T. Ono, T. Kawai, M. Tomizawa, and Y. Miyamoto. Demonstration of translucent elastic optical network based on virtualized elastic regenerator. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [63] S. K. Korotky, R. Essiambre, and R. W. Tkach. Expectations of optical network traffic gain afforded by bit rate adaptive transmission. *Bell Labs Technical Journal*, 14(4):285–295, February 2010.
- [64] J. Rahn, S. Kumar, M. Mitchell, R. Malendevich, H. Sun, K. Wu, P. Mertz, K. Crousore, H. Wang, M. Kato, V. Lal, P. Evans, D. Lambert, H. Tsai, P. Samra, B. Taylor, A. Nilsson, S. Grubb, R. Nagarajan, F. Kish, and D. Welch. 250gb/s real-time pic-based super-channel transmission over a gridless 6000km terrestrial link. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [65] D. J. Geisler, R. Proietti, Y. Yin, R. P. Scott, X. Cai, N. Fontaine, L. Paraschis, O. Gerstel, and S. J. Yoo. The first testbed demonstration of a flexible bandwidth network with a real-time adaptive control plane. In *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, September 2011.
- [66] N. Amaya, M. Irfan, G. Zervas, K. Banias, M. Garrich, I. Henning, D. Simeonidou, Y. R. Zhou, A. Lord, K. Smith, V. J. F. Rancano, S. Liu, P. Petropoulos, and D. J. Richardson. Gridless optical networking field trial: Flexible spectrum switching, defragmentation and transport of 10g/40g/100g/555g over 620-km field fiber. In *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, September 2011.
- [67] X. Cai, K. Wen, R. Proietti, Y. Yin, R. P. Scott, C. Qin, and S. J. Yoo. Experimental demonstration of adaptive combinational qot failure restoration in flexible bandwidth networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.
- [68] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, T. Yoshimatsu, T. Kobayashi, Y. Miyamoto, K. Yonenaga, A. Takada, O. Ishida, and S. Matsuoka. Demonstration of novel spectrum-efficient elastic optical path network with per-channel variable

- capacity of 40 gb/s to over 400 gb/s. In *Optical Communication, 2008. ECOC 2008. 34th European Conference on*, pages 1–2, September 2008.
- [69] D. J. Geisler, N. K. Fontaine, R. P. Scott, L. Paraschis, O. Gerstel, and S. J. B. Yoo. Flexible bandwidth arbitrary modulation format, coherent optical transmission system scalable to terahertz bw. In *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, September 2011.
- [70] K. Wen, Y. Yin, D. J. Geisler, S. Chang, and S. J. B. Yoo. Dynamic on-demand lightpath provisioning using spectral defragmentation in flexible bandwidth networks. In *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, pages 1–3, September 2011.
- [71] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos. Routing and spectrum allocation in ofdm-based optical networks with elastic bandwidth allocation. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–6, December 2010.
- [72] M. Jinno, H. Takara, and B. Kozicki. Dynamic optical mesh networks: Drivers, challenges and solutions for the future. In *Optical Communication, 2009. ECOC '09. 35th European Conference on*, pages 1–4, September 2009.
- [73] B. Kozicki, H. Takara, T. Yoshimatsu, K. Yonenaga, and M. Jinno. Filtering characteristics of highly-spectrum efficient spectrum-sliced elastic optical path (slice) network. In *Optical Fiber Communication - includes post deadline papers, 2009. OFC 2009. Conference on*, pages 1–3, March 2009.
- [74] X. Cao, V. Anand, Y. Xiong, and C. Qiao. A study of waveband switching with multilayer multigranular optical cross-connects. *Selected Areas in Communications, IEEE Journal on*, 21(7):1081–1095, September 2003.
- [75] R. Izmailov, Samrat Ganguly, V. Kleptsyn, and A. C. Varsou. Nonuniform waveband hierarchy in hybrid optical networks. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 1344–1354, March 2003.
- [76] A. N. Patel, P. N. Ji, J. P. Jue, and Ting Wang. Defragmentation of transparent flexible optical wdm (fwdm) networks. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2011.
- [77] Y. Yin, K. Wen, D. J. Geisler, R. Liu, and S. J. B. Yoo. Dynamic on-demand defragmentation in flexible bandwidth elastic optical networks. *Optics Express*, 20(2):1798–1804, January 2012.

- [78] P. R. Wilson, M. S. Johnstone, M. Neely, and D. Boles. Dynamic storage allocation: A survey and critical review. In *Proceedings of the International Workshop on Memory Management, IWMM '95*, pages 1–116, London, UK, UK, 1995. Springer-Verlag.
- [79] S. Mamagkakis, C. Baloukas, D. Atienza, F. Catthoor, D. Soudris, and A. Thanailakis. Reducing memory fragmentation in network applications with dynamic memory allocators optimized for performance. *Computer Communications*, 29(13-14):2612–2620, August 2006.
- [80] L. Alfaro and S. Gilmore, editors. *Process Algebra and Probabilistic Methods. Performance Modelling and Verification*, volume 2165. Springer Berlin Heidelberg, 2001.
- [81] W. J. Stewart. *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, 1994.
- [82] H. C. Tijms. *A First Course in Stochastic Models*. John Wiley & Sons Ltd, 2 edition, 2003.
- [83] Barrett et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. Philadelphia: Society for Industrial and Applied Mathematics, 1994.
- [84] W. Kahan. *Gauss-Seidel methods of solving large systems of linear equations*. PhD thesis, University of Toronto, 1958.
- [85] C. C. Paige and M. A. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Transaction on Mathematical Software*, 8(1):43–71, March 1982.
- [86] D. M. Young. *Iterative Solution of Large Linear Systems*. Academic Press, New York, 1971.