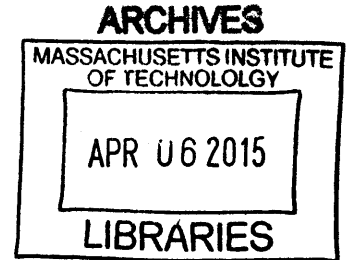# Network Management and Control of Flow-Switched Optical Networks:
## Joint Architecture Design and Analysis of Control Plane and Data Plane with Physical-Layer Impairments

by

Lei Zhang

B.Eng., Nanyang Technological University (2007)
S.M., Massachusetts Institute of Technology (2010)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

© Massachusetts Institute of Technology 2014. All rights reserved.

Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . **Signature redacted** . . . . . . . . . . . . . . . .
Department of Electrical Engineering and Computer Science
**Signature redacted** May 21, 2014

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Vincent W.S. Chan
Joan and Irwin Jacobs Professor of Electrical Engineering and
Computer Science
Thesis Supervisor

Accepted by . . . . . **Signature redacted** . . . . . . . . . . . . . . . .
\        𝒰 ◊   Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Students

# Network Management and Control of Flow-Switched Optical Networks:
# Joint Architecture Design and Analysis of Control Plane and Data Plane with Physical-Layer Impairments

by

Lei Zhang

Submitted to the Department of Electrical Engineering and Computer Science
on May 21, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

## Abstract

Optical Flow Switching (OFS) that employs agile end-to-end lightpath switching for users with large transactions has been shown to be cost-effective and energy-efficient. However, whether it is possible to coordinate lightpath switching and scheduling at a global scale on a per-session basis, and how the control plane and data plane performance correlate remained un-answered. In this thesis, we have addressed the network management and control aspect of OFS, and designed a network architecture enabling both a scalable control plane and an efficient data plane. We have given an overview of essential network management and control entities and functionalities. We focused on the scheduling problem of OFS because its processing power and generated control traffic increase with traffic demand, network size, and closely correlate with data network architecture, while other routine maintenance type of control plane functionalities contribute either a fixed amount or negligibly to the total efforts.

We considered two possible Wide Area Network architectures: meshed or tunneled, and developed a unified model for data plane performance to provide a common platform for the performance comparison of the control plane. The results showed that with aggregation of at least two wavelengths of traffic and allowing about two transactions per wavelength to be scheduled to the future, the tunneled architecture provides comparable data plane performance as the meshed architecture.

We have developed a framework to analyze the processing complexity and traffic of the control plane as functions of network architecture, and traffic demand. To guarantee lightpath quality in presence of physical-layer impairments, we developed models for quality of EDFA-amplified optical links and impairment-aware scheduling algorithms for two cases, a) the known worst case of channel quality is when there is no "On" channel in a fiber, and b) detailed channel configuration of a fiber is needed to determine channel quality. Without physical-layer impairments, tunneled architecture reduces control plane traffic and processing complexity by orders of magnitude. With impairment-aware scheduling, detailed channel configuration information report-

ing leads to heavy control traffic ($\sim$250 Gbps/edge); while known worst case and tunneling leads to manageable control traffic ($\sim$36 Gbps/edge) and processing power (1-4 i7 CPUs).

Thesis Supervisor: Vincent W.S. Chan
Title: Joan and Irwin Jacobs Professor of Electrical Engineering and Computer Science

# Acknowledgments

I own my greatest gratitude to my research advisor, Professor Vincent Chan. He has been supportive, patient and encouraging ever since my first day at MIT. He taught me formulating the right problem is just as important as getting the results, and the importance of always keeping an engineering intuition beyond mathematical equations or simulations. Over the years, he has also encouraged me to explore other aspects of life, and shared with me his insights, wisdom and his experience as a young MIT student. I am fortunate to have Professor Chan as my research advisor and life mentor. His passion, integrity, engineering intuition, and dedication have been, and will continue to be, my sources of inspiration.

I would also like to thank my thesis committee members, Eric Swanson and Professor Michael Watts. Eric's engineering wisdom, critiques, and suggestions have significantly improved this thesis, especially the models of quality of EDFA-amplified optical links. His passion in research and tenacity of drilling down to the very fundamentals have greatly inspired me, and will continue to be driving sources to me. I am also deeply grateful to Professor Watts for being patient with me and providing invaluable feedback for my research.

I have been fortunate to be a member of the "Chan-students", among which I have overlapped with Lillian Dai, Etty Lee, Guy Weichenberg, Andrew Puryear, Anurupa Ganguly, James Gettler, Mia Yinuo Qian, Mathew Carey, Henna Huang, Esther Jang, Shane Fink, Manishika Agaskar, Joseph Junio (post-doc), Jessica Weaver, Anny Xijia Zheng, and Antonia Allen. I am indebted to them for being a wonderful group of friends and enriching my Ph.D. life. I would also like to thank the administrative staff Donna Beaudry, and the visiting professor Angela Amphawan for their friendship and support.

MIT would have been less awesome without the amazing people I have met and made friends with. A sincere thank to all of those I have had the privilege to interact with. Special thanks to my karate buddies, Bonnie Lam, Ermin Wei, and Ying Yin, who have shared happy times and gone through difficult times (karate testings) with me, and always been there to "peer pressure" me. My gratitude also goes to my dear friends Cheng Ju, Qing He, Tong Wang, and Shen Shen for creating so many sweet and fun memories for me. In addition, I would like to thank my Japanese teacher, Masami Ikeda, for making every Japanese class an enjoyable time of language and culture.

Lastly, I would like to thank the most important people in my life, my family:

my parents for their sacrifices, unconditional love and support; my husband, Yehua Wei, for his love, understanding, support, and for being my best friend. This thesis is dedicated to them.

*Lei Zhang*
*Cambridge, Mass.*

# Contents

# List of Figures

13

14

15

# List of Tables

# List of Algorithms

# Chapter 1

# Introduction

Networks today are challenged by the ever-increasing growth of data traffic as we approach the "Zettabyte Era", when global IP traffic will reach 1.3 Zettabytes per year [2]. In the coming five years, the total Internet traffic is forseen to increase at a 23% compound annual growth rate (see Fig. 1-1). The Data traffic increase not only comes from the increasing number of users but more from new applications that entail large data transactions. For example, Internet video traffic (including Internet video to TV, video-on-demand, Internet, and P2P) amounted to 51% of all consumer Internet traffic in 2011, and will continue to increase to approximately 75% of total consumer traffic by 2017, as shown in Fig. 1-1. A large portion of these emerging large data traffic exhibits a bursty and unscheduled nature and has a heavy-tail distribution [5]. As illustrated by Fig. 1-2, with heavy-tail distribution, 1% of traffic sessions can contribute to 99% of the total traffic volume. This vast contrast with the traditionally assumed exponential-distribution of Internet traffic poses new challenges on network efficiency.

Another new trend of Internet traffic has been observed. From an analysis from Cisco [2], we are witnessing an explosion of data center traffic. As predicted, global data center traffic growth will increase three-fold by 2017, amounting to 7.7 Zettabytes annually by 2017. Global cloud traffic will grow faster than overall global data center traffic, and by 2017, will account for more than two-thirds of total global data center traffic. In addition, we are also witnessing a workload transition from local servers to the cloud. In 2012, 39% of workloads were processed in the cloud, and by 2017, nearly two-thirds of workloads will be processed by cloud data centers. Therefore, Internet is increasingly being used as a service bus for cloud-based services taking into users
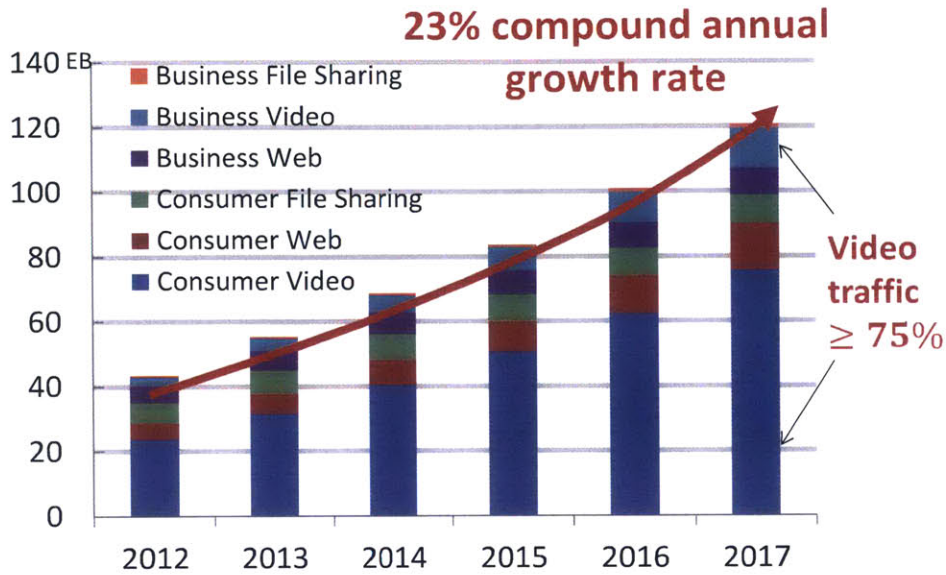
Figure 1-1: Internet traffic prediction from Cisco [2].

requests, and connecting the storage, data ingestion, and the processing units, Fig. 1-3. These new trends of the Internet usage could result in frequent data exchange through the Internet with requirements of low latency, high reliability, etc, and at the same time, call for an cost-effective and power-efficient transport.

This trend of fast increasing data traffic is soon going to out-pace the increase of the electronic processing speed of today's networks, which is dominated by Moore's Law. Optical networks are capable of providing bulk bandwidth that meets the traffic demand in the new trends in the form of optical transports. However, today's network is bottle-necked by electronic processing at routers or switches, and the current transport protocol limits an efficient use of the vast optical bandwidth. Thus, it calls for a new design of optical network architecture, in which along a connection optical-to-electrical-to-optical (OEO) conversion is abandoned at intermediate nodes and full potential of optical networks can be unleashed. Optical Flow Switching (OFS) [9] is an architecture construct that provides end-to-end all optical connections to users with very large transactions[1], and potentially can unleash the full potential of optical networks. OFS is a scheduled service in which connections are scheduled and set up prior to data transactions. Optical Cross Connects (OXC) and Wavelength-Selective Switches (WSS) are used at intermediate nodes to provide optical connections. This concept was proposed in 1988 by the "All-Optical-Network" (AON) Consortium [4, 11],

---

[1]Transaction of size of the order of Gigabytes that can hold the connection for hundreds of milliseconds or longer.

22
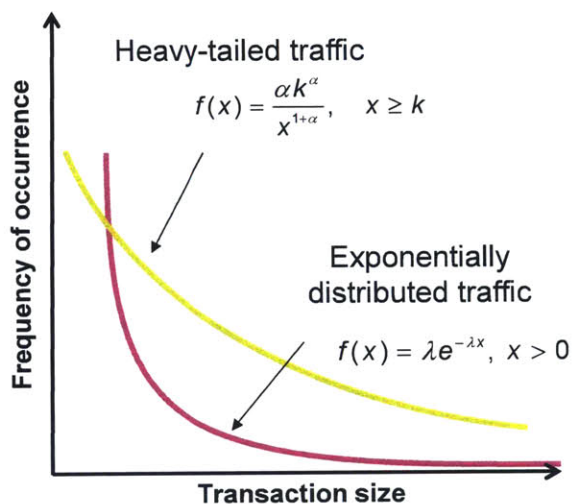
Figure 1-2: Illustration of the exponential probability distribution, and the heavy-tail distribution [5].

and was experimentally demonstrated in a network comprised of three nodes with optical switching capability at intermediate nodes [9]. It has been shown analytically that OFS is a cost-effective and energy-efficient transport service for large transactions [42, 26]. A viable physical layer design of OFS networks has been proposed and studied with present-day's matured technologies [42].

## 1.1 Related Work on Optical Flow Switching

Since 1988, various aspects of OFS have been studied [10, 22, 26, 41, 42, 47]. In this section we summarize related contributions to OFS.

### 1.1.1 Network Architecture and Flow Services

In [41], the authors proposed and studied a network architecture that simplifies the network management and control complexity. As shown in Figure 1-4, the architecture has the following attributes:

1. A Wide Area Network (WAN) with tunneled connections between WAN node pairs that changes slowly responding to traffic trend.

2. A reconfigurable Metropolitan Area Network (MAN) with an optimized physical topology. It responds to short term traffic demands and is switched at medium speed.

23
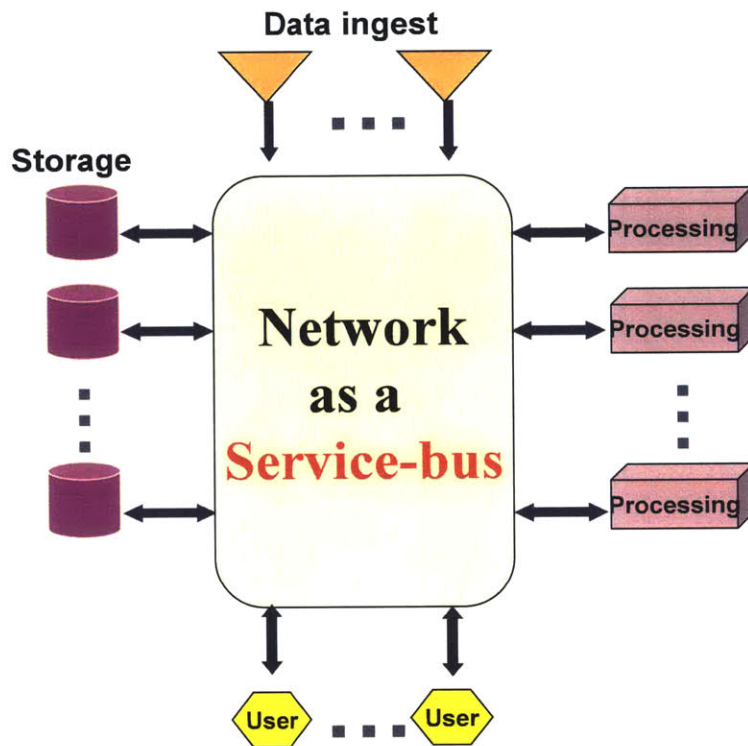
Figure 1-3: Illustration of the network as a service-bus.

3. A passive and broadcast Access Network (AN). It has an EDFA amplifiers chain pumped from the head-end [39] to allow more users to be aggregated into one AN, and thus statically share MAN/WAN resources.

**Topology of MAN and AN**

The MAN lightpath topology is based on generalized Moore Graphs, which minimize average hop count [19]. A generalized Moore Graph has a regular spanning tree that is fully populated except possibly the last level nodes. As shown in Figure 1-5, a MAN physical topology can be configured into a Moore graph lightpath topology using OXCs and patch-panels, and there exists a fully populated spanning tree of the MAN lightpath topology with the gateway node as the root. A splitter can be used at the gateway node to split the incoming light into all branches of the tree, allowing broadcast of the incoming data to all ANs attached to the tree. The construct of the embedded tree can be reconfigured at medium speed to accommodate short term traffic demands.

The AN topology also has a broadcast construct. As shown in Figure 1-6, Erbium-doped fiber amplifiers (EDFA) are pumped from the head-end of the passive optical
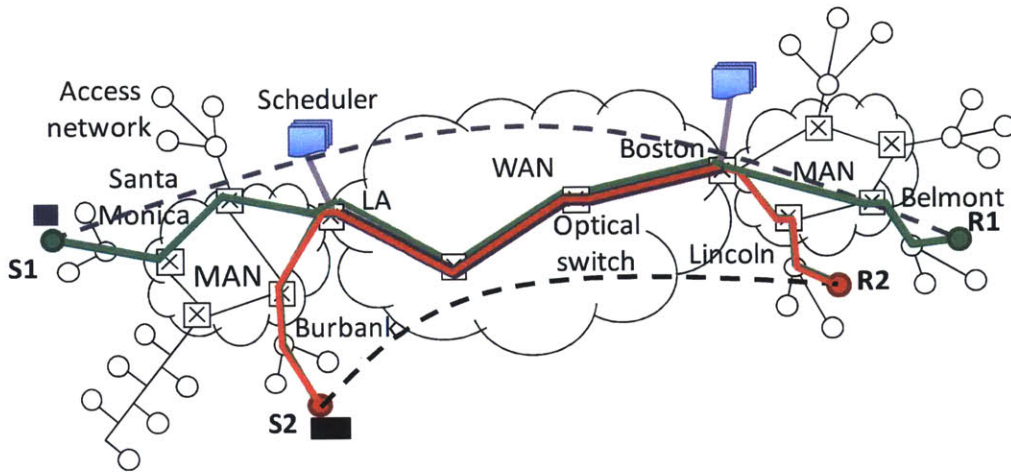
Figure 1-4: OFS architecture with WAN/MAN/AN.
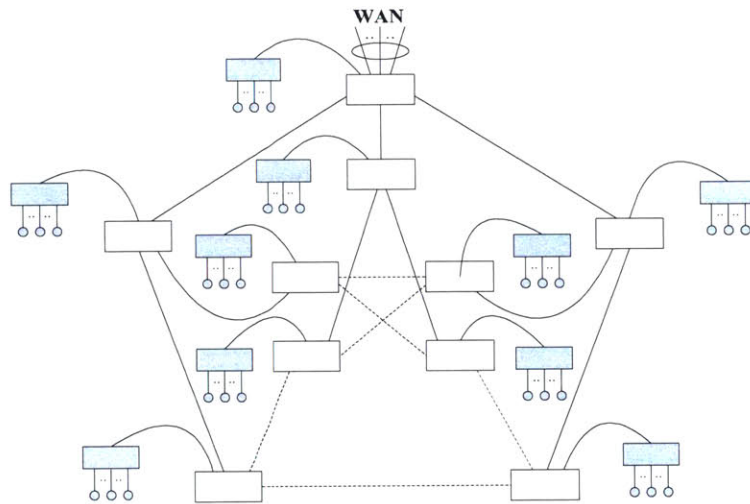


Figure 1-5: OFS MAN topology based on Moore Graphs. Reproduced from [9].

access network (PON). This allows more users to be statistically multiplexed in a broadcast network and, therefore, enables a more efficient use of MAN/WAN resources [39, 28, 8, 38].

**Scheduling**

With the above MAN/AN design, flows from a large number of users are multiplexed before entering the WAN. This leads to a statistical smoothing of the WAN traffic, rendering a quasi-static WAN logical topology applicable. This entails a substantial traffic decoupling between different WAN node pairs. To avoid collisions and to better utilize the WAN resources, a scheduling mechanism has been proposed and studied for
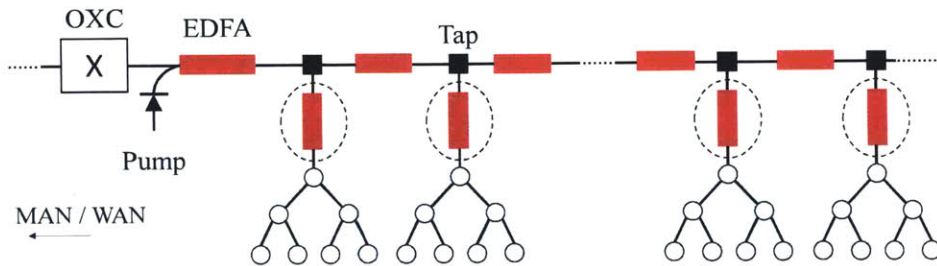
Figure 1-6: OFC access network physical topology: a remotely pumped broadcast passive optical access network. Reproduced from [39].

optical flow switching [40]. This scheduling algorithm reserves end-to-end lightpaths through schedulers upon end users' requests. In this algorithm, WAN/MAN wavelength channel is reserved first, followed by wavelength reservation in the source and destination ANs.

The following physical designs are adopted to make the scheduling algorithm simple.

1. Wavelength channels in the WAN are tunneled in a quasi-static manner. From the perspective of an individual flow, once it enters the WAN channel, it will only exit the WAN from its destined egress WAN node, and it sees no switching-in (or switching-out) of other flows into (out of) its channel.

2. In the MANs connecting the two ends of each WAN wavelength channel, there is a broadcast tree matched with the same wavelength from the headends. This means, once the WAN wavelength channel is available, the same wavelength will be available in the WAN, and the source and destination MANs.

3. Access networks are connected to the MAN through an optical switch.

With the above physical designs, contention of resources happens at two places, the WAN/MAN resources, and the AN resources. In the algorithm, two levels of queues, one at the WAN scheduler and one at the AN/MAN gateway, are used to resolve contentions.

The following example illustrates how the scheduling algorithm in [40] works. As shown in Figure 1-4, source node S1 in Santa Monica needs to send a large data file to destination node R1 in Belmont.

1. S1 sends its request for connection to the Scheduler at LA. The request enters the queue for the wavelength channels connecting LA and Boston.

2. When the request reaches to the head of the queue, LA Scheduler processes the request, assigns a wavelength channel, and informs the Boston Scheduler.

3. A secondary request for the same wavelength is sent to the secondary queues in Santa Monica and Belmont.

4. When the secondary requests reach the heads of the secondary queues and the same wavelength channel will become available in both source and destination ANs, the Schedulers in LA and Boston are notified.

5. The Schedulers in LA and Boston notify S1 and R1, respectively, about the scheduled time and connection for transmission. After the transmission, the resources are released from reservation for other users' use.

The time from the source node sends the transmission request to the scheduler, to the time an end-to-end connection has been setup, is the queuing/scheduling delay sees by the source node. Depending on the network loading, the size of the network, and the average transaction size this delay can be of the order of seconds.

## Probing

For some applications with urgent deadlines, the end user can pay more to bypass the normal scheduling process and set up an end-to-end connection via probing [10, 47], which only takes slightly more than the round-trip delay time.

In [10] and [47], the authors designed and analyzed the probing method for super-fast connection setups. As shown in Figure 1-7, in the probing process, the source node sends out probing signals along a set of independent paths and reserves the available resources along the probed paths; when the probing signals reach the destination node, the destination node picks a path that is available and reserved (if there are any) and notifies the source node of the selection; the destination node sends commands to the other reserved but unselected paths to release those resources from reservations.

To facilitate the source node to choose the set of paths to probe, in both [10] and [47], a control plane is used to periodically collect and broadcast network states. In [10], detailed information about each individual channel is collected, resulting in a huge control traffic. In [47], the information about the network states is collected and broadcast in the form of average entropy and mutual information evolution over time. The aggregation of information leads to reduced control traffic. An Entropy-Bellman-Ford (Entropy-BF) Algorithm has been designed to select path with lowest blocking

27

probability using broadcast entropy and mutual information and estimate the number of lightpath to probe to achieve a target blocking probability, in polynomial running time, resulting into significant reduction of complexity compared to the algorithm in [10].
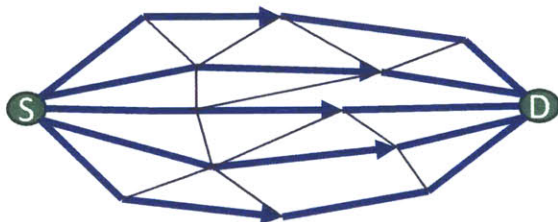
Figure 1-7: A schematic diagram of probing for OFS.

## 1.1.2 Physical Layer Impairments

In a long-haul fiber connection in today's optical networks, one common way of amplifying the optical signal before it degrades is to use EDFAs. Pumped by an out-of-band laser, an EDFA provides amplification to all wavelength channels in-band. In practice, there are normally two common operating modes for an EDFA: constant power mode and constant gain mode. Under the constant power mode, the control mechanism of an EDFA works to maintain the output power at a fixed level; while under the constant gain mode, an EDFA is controlled to keep the ratio of the output power to the input power constant. In long-haul networks, to mitigate the long-distance and avoid frequent regeneration of optical signals, a tandem of up to 40 EDFAs equally spaced are used for amplification of the optical signals. However, due to both the fundamental limitations and the constant gain control of EDFA, when lightpath is switched on and off dynamically in a meshed network, existing channels in the same fiber experience two types of impairments: fast transients, and steady-state channel quality variations. The former impairment can be quenched using adiabatic switching and a new Transport Layer Protocol as discussed in Section 1.1.3. The later impairments are results of a combination of possible causes including the randomness in the EDFA gain, accumulation of Amplified Spontaneous Noise (ASE), and the constant gain control system of EDFAs. The experimental results by Junio were published in [23]. Switching-induced channel-quality variations affect scheduling, i.e., worst case channel quality during the scheduled transmission time needs to be estimated to guarantee performance and reach (for a selected transmission rate if flexible transmission rate is used).

28

## 1.1.3 Transport Layer Protocol Design

In today's data network, Transport Control Protocol (TCP) in the Transport Layer is commonly used to provide reliable delivery of data and congestion control. Congestion control is carried out by slowly increasing the rate of releasing data packets into the network, and drastically reducing the rate when transmitted packets have not been acknowledged by the receiver after a certain amount of time. However, in optical flow switching, congestion is avoided by scheduling or probing to reserve the network resource, and a slow increasing of transmission rate of TCP is redundant and entails inefficient use of the high bandwidth of OFS connections. In [22], the authors designed and studied a new Transport Layer protocol that utilizes file segmentation and reassembly, forward error-correction (FEC), and frame retransmission.

The throughput and delay performance of OFS is studied for four example Transport Layer protocols: the Simple Transport Protocol (STP), the Simple Transport Protocol with Interleaving (STPI), the Transport Protocol with Framing (TPF) and the Transport Protocol with Framing and Interleaving (TPFI). From the analysis, the segmentation of a file into large frames (> 100 Mbits) combined with FEC, interleaving, and retransmission of erroneous frames (TPFI) is proposed for OFS.

## 1.1.4 Analysis of Cost and Energy Consumption

In [42, Chap.5], a performance-cost study is carried out to compare OFS to other prominent optical network architectures: Electrical Packet Switching (EPS), Optical Circuit Switching (OCS), Generalized Multiple Protocol Label Switching (GMPLS), and Optical Burst Switching (OBS). The results shows that the asymptotic normalized costs of competing architectures are several times higher than that of OFS, indicating that OFS is the most cost-scalable architecture among all the candidate architectures. Therefore, for sufficiently large average end-user data rates, OFS is a critical component for a hybrid network architecture and the most cost attractive homogeneous architecture.

In [26], the cost of optical networks has been studied with respect to power consumption. It is shown that, when quality of service and network flexibility, reliability and protection are considered, an optical bypass network is the most scalable with respect to power consumption. A hybrid design – in which whole wavelengths of core, stable traffic between node pairs are routed via direct, fixed lightpaths using patch

paneling, and bursty traffic is switched on a standard optical bypass network – further reduces the power consumption of a bypass network. Power distribution among different components is analyzed and the results indicate OXC switches are most scalable and O/E/O switches and routers are wasteful. Shortest path and minimum hop routing is proved to be power optimal and load-balanced routing should be avoided.

## 1.2  Challenges in Network Management and Control

With the related work discussed in the previous sections, what remains to be answered is whether coordinating (or scheduling) all user requests are implementable, in terms of control plane complexity and the volume of control traffic, while maintaining an efficient data network. In other words, a joint architecture of control plane and data plane needs to be studied to achieve high performance in both control plane and data plane.

OFS switched optical networks are very different from today's Internet. In an OFS network, end-to-end all optical connections are realized by replacing routers with optical switches at intermediate nodes. This substitution of routers with optical switches entails the following critical changes:

1. Packet routing capability is lost. In today's metro and wide-area networks, lightpaths are terminated at routers/switches at intermediate nodes where routing/switching is performed based on headers of data packets, and at the output port, data is regenerated into optical signals and sent along lightpaths. However, in OFS networks, because of the limitations of optical signal processing and buffering technologies, it is hard to carry out on-the-fly routing/switching directly on optical signals at optical switches. Instead, the right mapping of input to output ports is carried out in the connection-setup process before data transmission.

2. Visibility into the optical signal is lost. In traditional data networks where OEO conversions are carried out at routers/switches, visibility into the optical signal comes for free, enabling easy access to monitoring of many network management metrics, for example, signal-to-noise ratio (SNR), loading factor, etc.

30

With the above two critical changes, OFS networks cannot be managed in the same way as today's data networks, and Network Management and Control (NMC) concepts need to be reconsidered. Without routing at intermediate nodes, to set up end-to-end connections, some network entities need to perform path computation, which then requires knowing the network topology and network state. Scheduling of connections needs to be done to avoid contentions. Without visibility and the free monitoring capabilities, the OFS networks elements need to be monitored deliberately. In addition, discussed in [9], the length of OFS sessions can be as short as one second. Therefore, in an OFS network there will be frequent transaction requests that have to be handled by the Network Management and Control (NMC) system. Thus, we need to reconsider the Network Management and Control concepts for OFS networks and a redesign of the control plane is necessary.

In OFS networks, an off-band control plane is used to carry network management and control traffic. Residing on the control plane, dedicated NMC entities (including Network Manager, Scheduler and Element Control Units) are used to perform network management and control functions. Network manager (NM) is the control unit that oversees the whole network, and performs NMC functions that are of a global scale and favor centralized control, e.g., maintaining the network topology and network states from updates from the other NMC entities, performing end-to-end path computation and load balancing, fault diagnostic, etc. A Scheduler performs scheduling and coordinating of end-to-end connection setup upon user's request. Element Control Units (ECU) monitor operating states of active network components, periodically send updates to the NM, and execute commands from the NM.

To manage and control a network that is of such a large scale, we need to study the scalability of our NMC algorithms and computations, and the corresponding sensing requirement. We also need to study the amount of control traffic that passes through the control plane, as high-volume control traffic can yield a non-scalable control plane. However, these NMC performance factors, do not only depend on the design of the NMC system itself, but is also related to the architecture design of the flow-switched optical data plane.

# 1.3   Data Plane Architecture

The data network spans across the WAN, MAN and Access Network. To manage and control such a large scale network, scalability of control plane function computations and amount of control traffic need to be carefully investigated. However, since the characteristics of the data plane architecture (in terms of size, topology, etc) are inputs to the NMC functions, NMC scalability and performance depends on the architecture of the data plane, through which performances of data plane and control plane are coupled.

Because the WAN resources are most expensive and should have high utilization, we focus our design and analysis in the WAN in this thesis[2]. [48] shows that when users can tolerate delay of at least one transaction duration, with traffic aggregation of two wavelengths into one WAN node, a quasi-static MAN and Access Network architecture is preferred. Therefore, the amount of control traffic and computation complexity for scheduling coordinations in the MAN and AN is negligible compared to those in the WAN. Therefore, limiting our analysis in the WAN does not disqualify the significance of our analysis.

One natural architecture design is to allow full switchability at all nodes in the WAN, providing possibly maximum data network capacity but exacerbating NMC complexity. We call this mesh architecture with full switchability, Architecture M. Another architecture arises from the data aggregation effect of the hierarchical physical topology, in which the WAN serves as the backbone for transport of huge data, and the MANs cover large areas and serve as data collection points for the WAN. Although data generated from end users may be bursty, after data aggregation at the ANs and MANs, heavy and smooth traffic with a quasi-static pattern enters the WAN. Therefore, it is reasonable to assume a quasi-static tunneled logical topological architecture for the WAN. We call this Architecture T. In this thesis, we will use Architecture M and Architecture T as example architectures to analyze the dependence of control plane and data plane performance on network architecture.

---

[2]Readers who are interested in architecture design and performance analysis of the MAN and AN can read [48] and [39].

## 1.4 OFS vs. Optical Circuit Switching and GMPLS

In the field of all-optical architectures, people often confuse optical flow switching with optical circuit switching (OCS), and the optical flow switching network management and control with Generalized Multi-Protocol Label Switching (GMPLS) [27]. Next we will discuss differences between OFS and OCS, and differences between OFS network management control with GMPLS. Readers who are familiar with them could skip this section and continue with Section 1.5.

### 1.4.1 OFS vs. OCS

Optical Flow Switching (OFS) is a network architecture that enables per-transaction based lightpath circuit switching from end user to end user, with architecture construct from the physical layer to the Transport Layer and from the WAN, MAN to the AN. It is a design of both the control plane and the data plane, and this specific architecture is designed to be scalable and manageable, even when ultra-fast (lightpath holding time can be as short as one second) end-to-end lightpath switching is provided with low blocking probabilities in a highly loaded network. The architecture concepts germinated 26 years ago and have been actively researched since then. Optical Circuit Switching also enables lightpath circuit switching in all-optical networks. However, Optical Circuit Switching has been mostly studied for the core network, and once a circuit is set up, it typically lasts for days, and usually numerous transactions go through the circuit before it is torn down. Conceptually, the concepts of Circuit Switching might be extended from the core network to the MAN and AN, and to provide per-flow connection setups and tear-downs, but the problems of control plane scalability and data plane performance (low blocking probability at high utilization and fast setup time ~one round-trip propagation time) have not been addressed by OCS and will need significant architecture development.

### 1.4.2 OFS vs. GMPLS

Generalized Multi-Protocol Label Switching (GMPLS) is a protocol suite extending Multi-protocol Label Switching (MPLS) [1], which directs data from one network node to the next based on short path labels, avoiding complex lookups in a routing table. GMPLS is the de facto control plane of Wavelength Switched Optical Network. Optical Flow Switching came before GMPLS, and is different from GMPLS in the following

33

ways:

1. Scheduled end-user to end-user connections (traversing the AN, MAN and WAN) in OFS is per-transaction based (a connection is set up for the source node to send data to the destination over one transaction, and the connection is torn down immediately after the transaction ends). The connections are scheduled into the future if there are no available network resources at the time of request. Scheduling into the future allows high network throughput to be achieved. GM-PLS can be used among routers to set up lightpath circuits, typically in the core network, and those circuits are not per-transaction based and usually last for more than one transaction, sometimes with idle frames between real traffic. GMPLS does not support scheduling into the future time horizon if a request for a circuit cannot be satisfied at the time of request.

2. Even when providing per-transaction based connection (connection can be as short as one second) set-ups and tear-downs, the architecture of the OFS control plane is designed to be scalable and manageable while achieving low blocking probability in networks of medium to high loading (a result of this thesis). GM-PLS is capable of setting up and tearing down circuits at a slower time scale and the circuits last much longer than one second. If GMPLS is used to do fast circuit switching, it will be challenged by the huge amount of control traffic necessary to update the network state, and it will be hard to achieve fast setup times with low blocking probability and reasonable network loading without significant architecture development. In fact, these challenges haven't been addressed by the GMPLS community.

3. The Entropy-Bellman-Ford Algorithm of OFS described in [47] is for a special operation mode of OFS which allows end users with privileged access to the network (financial sector premium service, for example) to have instant services (a fast setup time ~one round-trip propagation time) instead of the normal scheduling operation mode of OFS. The Entropy-BF Algorithm proactively probes the network along multiple pre-selected paths to find an open end-to-end connection with a delay time of one round-trip with low blocking probability even in highly loaded networks. This ultra-fast connection setup mechanism is not provided by the GMPLS. Even if GMPLS is directly generalized to do probing, it will be difficult to guarantee low blocking probability in a network with super-fast network

34

state dynamics without over-probing which may lead to low network utilization. This is because, even with super-fast updates of network states which may overwhelm the control plane, the network state information is going to be stale when it is sent to every node in the network because of network delay (propagation delay at least). To achieve low blocking probability, the Entropy-BF algorithm uses probabilistic parameters to capture the evolution of network state and broadcast the evolution to each node so that the network state can be predicted from the stale network state broadcast. In addition, compressed (not detailed) network state evolution is collected and disseminated to avoid flooding the control plane with huge control traffic. A user can use the broadcast information to predict the network state and choose multiple paths over lightly loaded network regions to probe and reserve available network resources, so that a connection can be set up with a low blocking probability.

## 1.5 Thesis Scope and Organization

This thesis focuses on the network management and control for all-optical flow switched networks, from perspectives of control complexity, control traffic, dependence of network management and control efforts on data network architecture, and the impact on data plane performance. Two example data network architectures, meshed or tunneled, are used in the analysis. The presence of physical layer impairments (as discussed in Section 1.1.2) resulted from fast lightpath switching creates more challenges: (a) lightpath quality might need to be monitored to guarantee performance, and (b) the scheduling algorithm needs to make sure the assigned lightpath satisfies a minimum channel quality requirement. Therefore, we have also developed models for the quality (in terms of Bit Error Rate) of EDFA-amplified optical links, designed scheduling algorithms using these models, and analyzed the required processing power and control traffic volume. The rest of the thesis is organized as follows:

Chapter 2 gives an overview of the control plane, introduces the control plane components, and analyzes the corresponding control functions from perspectives of time scales and input size.

Chapter 3 develops a unified model to analyze the scheduling performance (blocking probability, throughput and delay) as a function of network topology, traffic demand, network resources, and compares the performances of Architecture T and M.

Chapter 4 studies the network management and control efforts generated by the scheduling of optical flow switching. A brief literature on related work in routing and wavelength assignment and OFS scheduling is presented. The OFS scheduling problem is decomposed into a routing and scheduled wavelength assignment problem. Algorithms are then designed for each problems, and analyzed in terms of processing power and control traffic.

Chapter 5 examines the physical layer impairments in EDFA-Amplified optical networks. We have developed models for channel quality of EDFA-amplified optical links, and compared our models with experimental data in [23].

Chapter 6 re-examines the network management and control taking into account of the physical layer impairments. Impairment-aware routing and scheduled wavelength assignment algorithms have been designed and analyzed.

Chapter 7 concludes the thesis.

# Chapter 2

# Introduction to OFS Network Management and Control

In flow-switched optical networks, data is sent end-to-end with optical signals in the data plane, in the form of optical flows, bypassing routers and O-E-O conversions. To accommodate the stochastic nature of transmission requests, all optical end-to-end connections will be set up and torn down dynamically, and a control plane is used to manage the whole network, perform scheduling, network reconfiguration, and faults diagnostic, etc. However, the large scale of the network and the stochastic nature of the connection requests pose huge challenges on designing a scalable, efficient, and easy-to-implement control plane without compromising too much of the network throughput. The network management and control efforts highly depend on the design of the network architecture. Our work intends to provide insights into Network Management and Control for optical flow switched networks and its correlation with network architecture designs. We propose and examine the necessary Network Management and Control functions, analyze their computation complexity and scalability, and study the amount of control traffic with respect to different architectures.

In optical flow-switched optical networks, the substitution of routers with optical switches entails the following changes:

1. In OFS networks, user end-to-end transmissions are scheduled to avoid collisions.

2. Without O-E-O conversions, visibility into the optical signal is not as easily accessible and OFS networks need to be monitored by tapping, detecting and analyzing the optical signals.

In addition, the length of OFS sessions can be as short as one second, [9], and hence there will be frequent transaction requests that have to be handled by the Network Management and Control system, and a complete redesign is necessary.

## 2.1 Control Plane Overview

Besides providing basic NMC functionalities including network monitoring, scheduling and connection setups, fault diagnostics, etc, a network management and control system should be future-proof and failure tolerant. In other words, the NMC system should be self-adaptive to adjust to traffic increases and unpredictable traffic pattern changes, and when failure occurs, should still be able to manage and control the networks and take measures to ease the situation.
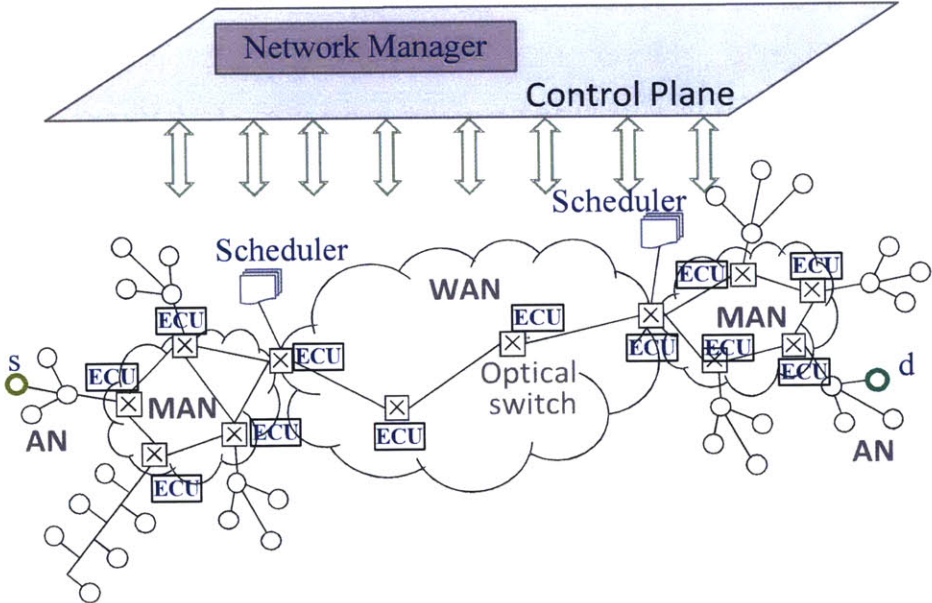


Figure 2-1: Illustrations of control plane and data plane, and NMC entities with distributed scheduling. ECU is short for Element Control Unit.

As marked in blue in Fig. 2-1, three NMC entities participate in the network management and control: Network Manager (NM), Scheduler (Sch), and Element Control Unit (ECU). Each network entity participates in certain control functions. NM resides in the control plane, manages and controls the whole data network, and performs centralized NMC functions, including path computation, fault diagnostics, and load balancing. It also carries out tunnel setups and tear-downs, collects link state updates from Schedulers and ECUs, and responds to end user when end user detects high

BER from the received signal. A Scheduler performs scheduling and coordinating of end-to-end connection setup upon user's request. There can be one Scheduler for the whole network (for centralized scheduling) or one Scheduler at each WAN node (for distributed scheduling). Scheduler maintains queues of requests and transmission sequences, computes schedules, monitors traffic statistics, reports to NM, requests tunnel setups and tear-downs when necessary, and notifies source and destination nodes of the schedule and designated lightpath connection. An ECU resides at each active network component, monitors status of network components (EDFA, OXC, etc), detects failure, reports to NM, and acknowledges commands of reconfiguration from NM. An end user requests a schedule and connection, monitors channel BER when receiving data, reports to NM if channel BER is high for decoding, and when having urgent data to send, probes to set up connections.

Below we list network management and control functions for the three network entities and actions from end users, the interactions among which are summarized in Fig. 2-2.

1. Network Manager

    - Path computation

    - Fault diagnostics

    - Sends commands to set-up/tear-down tunnels

    - Collects link state updates from Schedulers, ECUs. Responds to end user when end user detects high BER from the received signal

    - Perform load balancing

2. Scheduler

    - Maintains queues of requests and transmission sequences

    - Computes schedules

    - Monitors traffic statistics

    - Reports to NM (link state update)

    - Requests tunnel set-up/tear-down

    - Notifies source and destination of the schedule and designated lightpath connection

39

3. Element Control Unit

- Monitors equipments operating modes (EDFA, OXC, etc)

- Monitors channel qualities (for Impairment-aware scheduling)

- Report channel qualities to Sch and NM (for Impairment-aware scheduling)

- Reports to Network Manager (regular link state update other than impairments reporting)

- Detects failure

- Acknowledges commands of reconfiguration from NM

4. User

- Requests schedule and connection

- Monitors channel BER when receiving data

- Reports to NM if channel BER is high for decoding

- Probes to set up connections
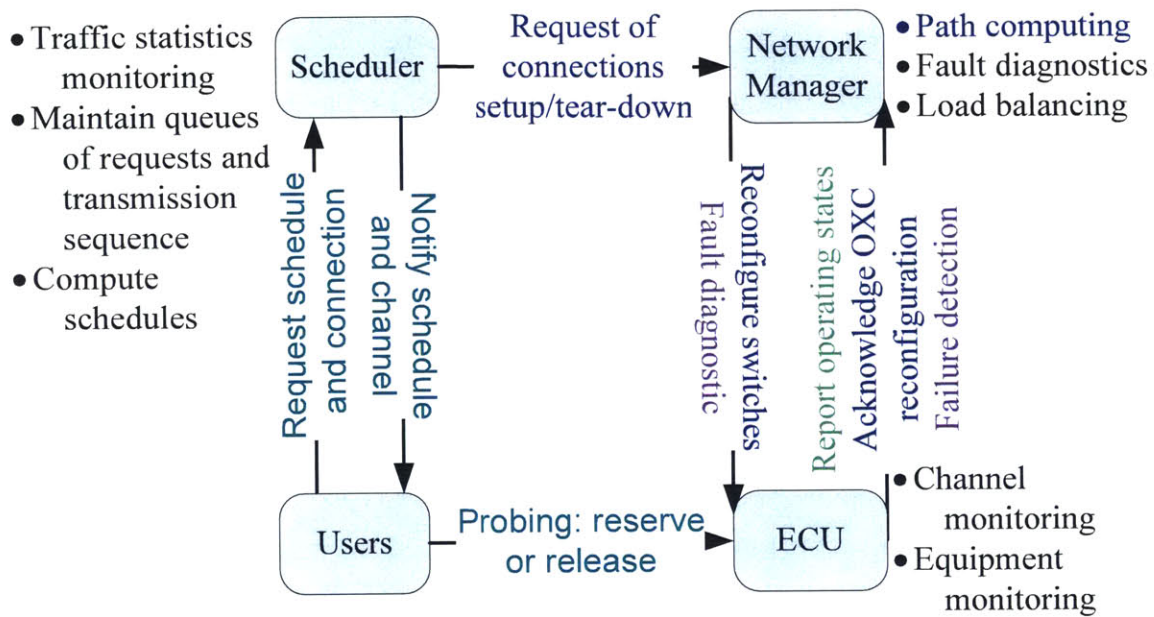


Figure 2-2: Interactions of NMC entities.

## 2.2 NMC Functions

When a source node (end user) needs to send a large file to a destination node, it sends a request either to the centralized scheduler (for centralized scheduling) or the Ingress Scheduler at the WAN node the AN of the source node is connected to (for distributed scheduling). At the Scheduler, the request is first put into a first-in first-out (FIFO) queue, and when the request moves to the head of the queue, the Scheduler computes a schedule and an end-to-end connection, notifies the result to the source and destination nodes, and then puts the request into a transmission sequence, so that the Scheduler can work on the next request in the FIFO queue. Both the queue and the transmission sequence are maintained by the Scheduler. The corresponding control functions are called at every session, and are per-session based.

Recall that in Section 1.3, we have defined two data network architecture for the WAN, meshed (Architecture M) or tunneled (Architecture T). Architecture M allows full switchability at all nodes in the WAN, providing possible maximum data network capacity but exacerbating NMC complexity, while Architecture T is a quasi-static tunneled logical topological architecture for the WAN. In Architecture M, with full switchability, scheduling is carried over the global mesh network, and tracks the traffic statistics from the past to the future booking (with a limited horizon) of the whole network. These network statistics are shared with the Network Manager, and are useful for network planning, that is, if the average loading of some network region exceeds a certain threshold, the Network Manager can deploy more network resources for that region, and if the average loading of a network region drops below a certain threshold, the Network Manger can reassign a portion of the network resource of that region for other use. Since average loading is used as the indicator, such network resource deployment and reassignment is expected to occur at a much slower time scale than the per-session time scale. By slowing down the resource allocation speed in the MAN/WAN this architecture feature is one important step towards NMC scalability.

In Architecture T, the connections in the WAN are tunneled. When the average loading of a wavelength tunnel exceeds a certain threshold, the Scheduler can request the Network Manager to set up a new tunnel to serve the same WAN node pair. Similarly, when the average loading of the tunnels serving the same WAN node pair drops below a certain threshold, the Scheduler can request the NM to tear-down one wavelength tunnel. From our assumption that the aggregated traffic across the WAN is quasi-static, we expect the control function of Connection Setup and Tear-down be

called only quasi-statically. On the NM side, when it receives such requests from a Scheduler, it will carry out computation locally, deciding which the affected ECUs are. The NM then commands the ECUs to carry out reconfigurations. Subsequently, this control function is also quasi-static.

A Scheduler monitors the traffic statistics of a subset of the network in Architecture T, or the whole network in Architecture M. ECUs monitor locally the operating status of the network elements. Receiving end user monitors the BER constantly when it receives the data. Schedulers and ECUs periodically summarize the states of the elements they monitor and report to the NM. Since we expect most of the active network elements not to change operating states frequently (except for switches in Architecture M), the periodic update can be run at a large time scale between updates, but ECUs run link state updates if they see a change in the operating state or an error. An end user only signals the NM when it detects a high BER, which happens occasionally and is event-driven. When an NM receives Error messages from ECUs or User, it performs fault diagnostics over all the possibly involved network elements. With the presence of switching-induced physical layer impairments, whenever a transmission starts or ends the channel configurations of involved fibers change, leading to link impairment reporting at a per-session time scale.

Based on the above discussion, we summarize the NMC functions in Table 2.2, and map the functions onto a 2-D graph of time scale and number of involved nodes in Fig. 2-3. In Table 2.1, we list the participants of network management and control functions, with entry "✓" indicating the function contributes to control traffic, and entry "o" indicating the function runs locally and does not generate control traffic.

From Fig. 2-3, and Tables 2.2 and 2.1, we observe that there are two classes of control functions, one fast and one much slower:

1. The fast (<1s), scheduling function which is per transaction and contributes most to the the total NMC computation complexity and control traffic at the requests arrival rate (for both Architecture M and Architecture T). For impairment-aware scheduling, channel quality is monitored and updated on a per-session base and also contributes to large control traffic volume.

2. Slower link state updates, e.g., fault diagnosis and load balancing in the MAN and WAN.

Therefore, to simplify analysis, we focus on the scheduling part of NMC, first the

generic scheduling, and then the impairment-aware scheduling. Since the WAN resources are the most costly and have the highest energy consumption it should have high utilization. Thus our analysis will be limited to the wide area backbone network.

Table 2.1: Participants of control plane functions, with NM for Network Manager, Sch for Scheduler, and ECU for Element Control Unit. Entry "✓" means the function contributes to control traffic, while entry "o" means the function is carried out locally.

| Control Plane Functions | | NM | Sch | ECU | User |
|---|---|---|---|---|---|
| **Scheduling** | Request/notify | | ✓ | | ✓ |
| | Connection Setup | | ✓ | ✓ | |
| | Compute | | o | | |
| **Tunnel Setup and Tear-down** | Request/notify | ✓ | ✓ | | |
| | Compute | o | | | |
| | Command/ACK | ✓ | | ✓ | |
| **Monitoring** | Traffic statistics | | o | | |
| | Equipment | | | o | |
| | Channel | | | | o |
| **Link State Update** | Periodic | ✓ | ✓ | ✓ | |
| | Switching-driven channel quality change | ✓ | ✓ | | |
| | Error-driven | ✓ | ✓ | ✓ | ✓ |
| **Load Balancing** | Compute | o | | | |
| | Reconfigure | ✓ | ✓ | ✓ | |
| **Queue and Transmission Sequence** | | | o | | |
| **Fault Diagnostics** | | ✓ | | ✓ | |
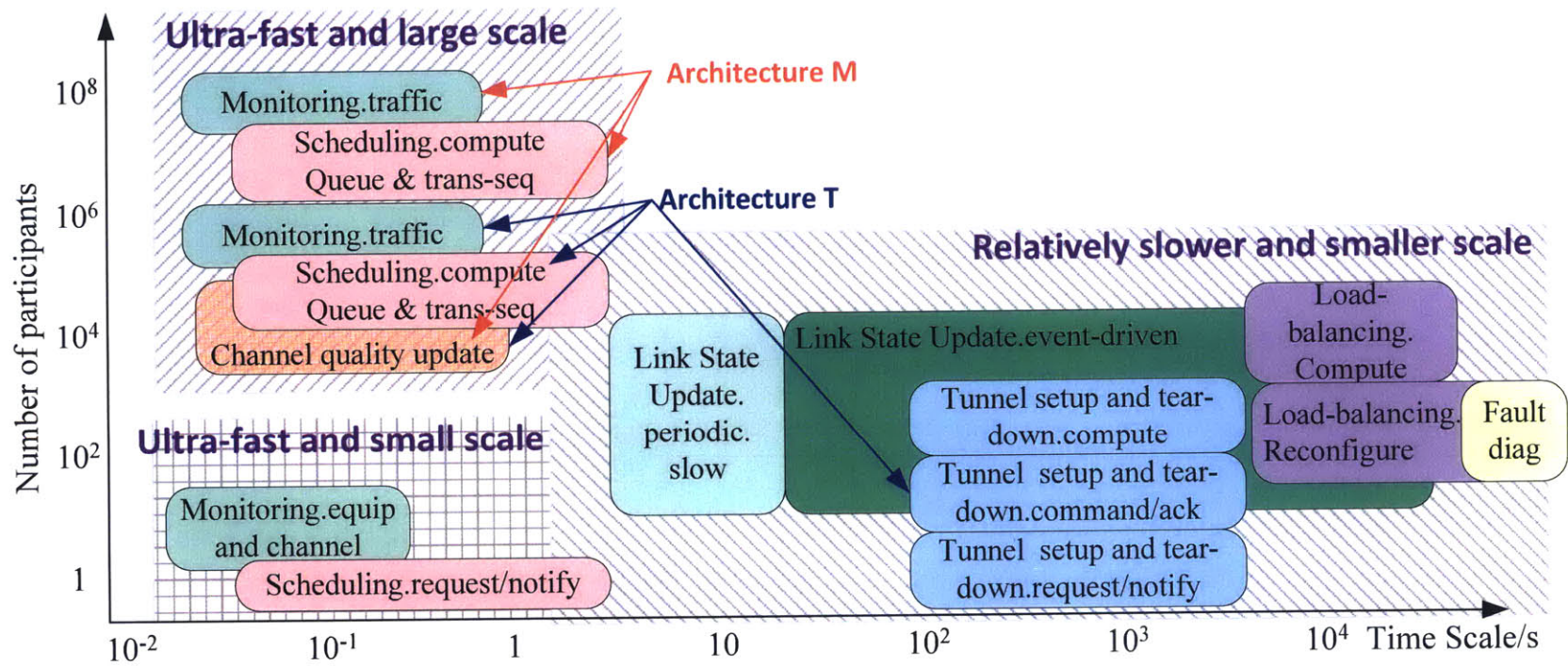
Figure 2-3: Mapping of control functions to time scale and number of nodes for Architecture M and T. Functions unspecified with an architecture apply to both architectures.

Table 2.2: Control plane functions of different time scales.

| Control Entities | Local Functions | Network control functions that contribute to control traffic | | | |
|---|---|---|---|---|---|
| | | Super fast | Quasi-static | Periodic | Occasional |
| Network Manager (NM) | 1. Path computation<br>2. Load balancing | | Command **ECUs** to set up or tear down connections | | 1. Fault diagnostic through **ECUs**<br>2. Commands **ECUs** to reconfigure for load-balancing |
| Scheduler | 1. Maintain queues of requests and transmission sequence<br>2. Compute schedule<br>3. Monitor traffic statistics | Notify source and destination nodes of the schedule and connection | Request **NM** for tunnel setup\tear-down | Report traffic statistics to **NM** | |
| Element Control Unit (ECU) | 1. Monitor equipments operating modes<br>2. Failure detection<br>3. Monitors channel qualities | Report channel qualities to Sch and NM | Acknowledge **NM** for proper or improper reconfigurations of OXCs | Report to **NM** | Report to **NM** upon failure detection |
| User | Monitor channel BER | 1. Request **Sch** for schedule and connection<br>2. Probe **ECUs** to reserve or release resources | | | Report to **NM** if channel BER is high for decoding |

# Chapter 3

# Data Plane Performance Analysis

Data plane performance is coupled with control plane performance through network topology and network control algorithms (routing, scheduling, etc). It also depends on the network resources, and network loading (and thus, traffic demand). Assume wavelength continuity constraint is respected by the scheduling algorithm, because the current wavelength converting technology is not cost-efficient and we don't foresee it to be mature soon. In [7], Barry and Humblet studied blocking probability of unscheduled circuit-switched all-optical networks with and without wavelength converters. We name their model the Barry-Humblet Model. They introduced three separate models, investigating the effect of path length, switch size, and interference length, respectively, on blocking probabilities. In their models, a session request between A and B is blocked if, for all-optical network without wavelength converters, all lightpaths (each lighpath is composed of channels of the same wavelength) are blocked, or, for all-optical network with wavelength converters, there exists a hop with all wavelengths used. However, the three parameters they investigated separately are all manifestations of the network topology and lightpath switching mechanism, and are correlated, and the Barry-Humblet Model did not give an analysis of the overall effect of network topology on blocking probability. In this Chapter, we will develop a unified model to investigate the network topology on blocking probability of circuit-switched all-optical networks with no wavelength converters, and extend it to flow-switched optical networks with scheduling.

47

## 3.1 OFS Scheduling with Schedule holders

As illustrated in Fig. 3-1, when a request arrives, the Scheduler will first try to accommodate it with network resources that are available at the time of the request along a candidate path[1]. If there are no available resources at that time, the Scheduler schedules the session to a lightpath with the first available time in the future. A scheduled transaction in the future is stored by "schedule holders" along the link[2] between adjacent node pairs. In the case of the tunneled architecture, the schedule holders are located at the Ingress WAN node only. A schedule holder can be considered as a marker to the corresponding wavelength channel specifying the future assignment and time of usage of this wavelength channel, in the form of $\langle$wavelength, $t_{\text{start}}, t_{\text{end}}\rangle$. It has two states, "available" or "occupied". Only available schedule holder can be assigned with new transactions, and upon assignment its state changes from "available" to "occupied". If there are only finite $N_s$ schedule holders for each link, sessions will not be scheduled into the infinite future. Assume there are $\Lambda$ wavelength channels for each link. Define the scheduling capacity to be the ratio of $N_s/\Lambda$. A lightpath is available for a certain time duration in the future if there is at least one schedule holder that is available for each hop along the lightpath throughout that duration. A request is blocked if there are no currently available network resources and there are no available lightpaths in the future.

We consider network topologies that are symmetric or close to symmetric, and assume the network is in steady state in the following analysis. Define $\rho$ to be the network link loading. Since the network is in steady state, $\rho$ is also the probability that the link is occupied. Define $P_n$ as the probability that a session enters the link, and $P_l$ as the probability that an existing session leaves the link (Fig. 3-2). Use a two-state ("available" or "occupied") Markov Chain (Fig. 3-3) to model the state of the link[3]. Then the transition probability from state "available" to "occupied" is $P_n$, and the transition probability from the state "occupied" to the state "available" is $P_l(1 - P_n)$. Therefore,

$$\rho = \frac{P_n}{P_n + P_l - P_n P_l} \tag{3.1}$$

---

[1]The path here refers to a path in the logical network topology.

[2]A link refers to an edge connecting two nodes in the logical network topology. A link in Architecture M corresponds to all wavelength channels connecting a node pair which are only one-hop away in the physical topology. In Architecture T, a link corresponds to the tunneled lightpaths connecting a node pair (which can be one-hop or multi-hop away in the physical topology).

[3]The state of a link is usually modeled as a Markov Process, which can be approximated using a sampled Markov Chain.
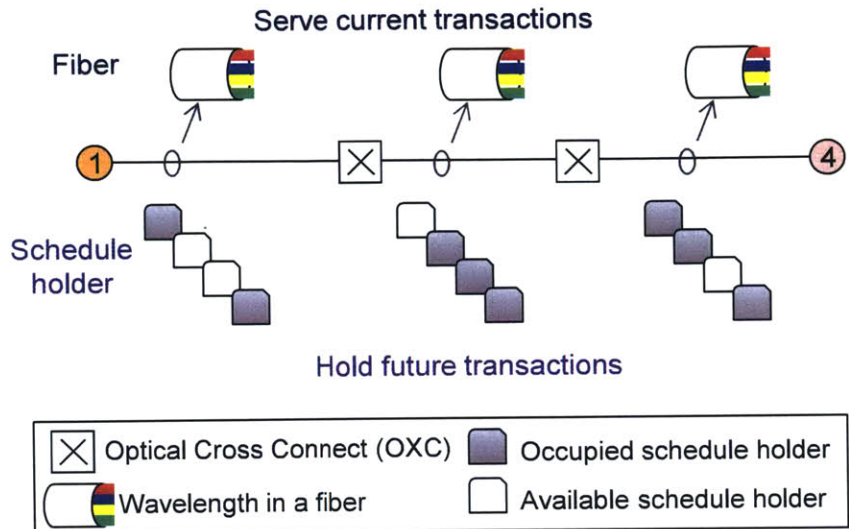
Figure 3-1: Illustration of OFS scheduling with schedule holders.
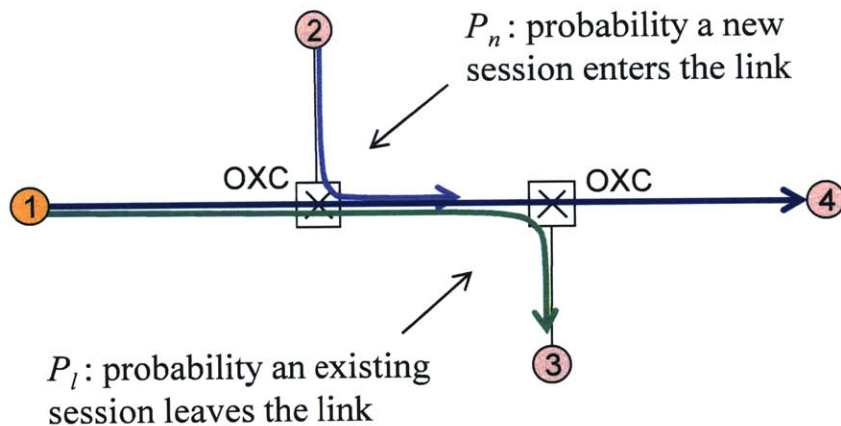


Figure 3-2: Illustration of traffic merging into and diverging from a path.

## 3.2 Blocking Probability for Architecture M

Consider a mesh network graph with $N_V$ nodes, $N_E$ edges, and $\Delta$ average node degree. At one intermediate node along a lightpath, define $P_{\lambda l}$ as the probability that an transmission leaves this lightpath, and $P_{\lambda n}$ as the probability that a new transaction enters this lightpath[4]. To derive $P_{\lambda l}$, consider a link $L_{ij}$ connecting node $i$ to node $j$ on a designated path. Define $N_{th}$ as the number of lightpaths that go through $L_{ij}$, and $N_{tm}$ as the number of lightpaths that enter $L_{ij}$ at node $i$ and terminate at node $j$. Assume a uniform all-to-all traffic model and that there is one unit of traffic from one

---

[4]$P_{\lambda n}$ and $P_{\lambda l}$ are special cases of $P_n$ and $P_l$, respectively.
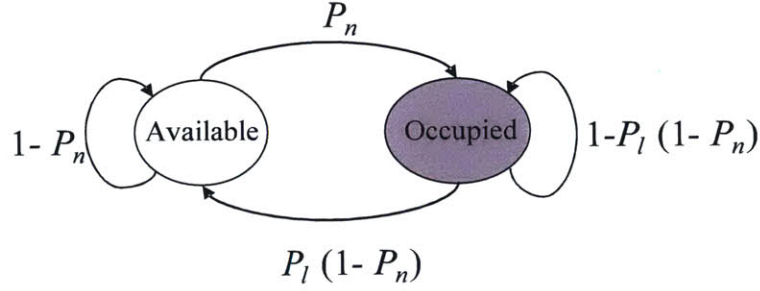
Figure 3-3: Markov Chain model of link state.
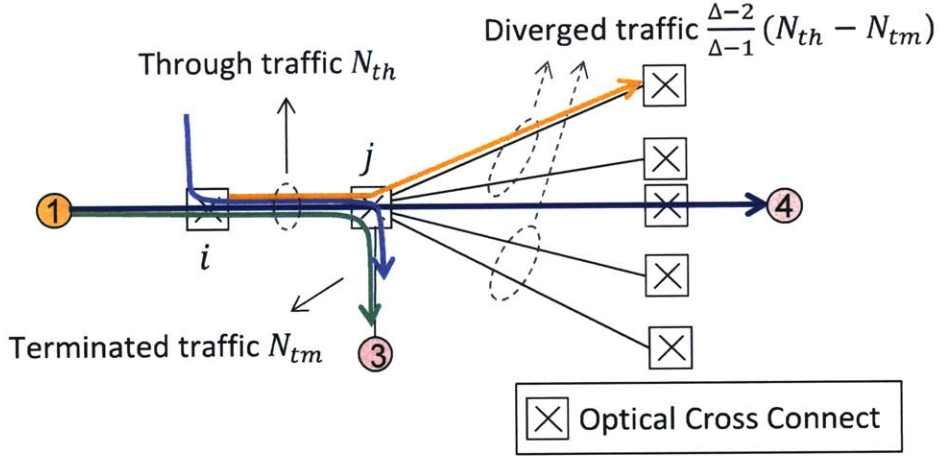


Figure 3-4: Illustration of lightpaths going through link $L_{ij}$, lightpaths terminated at node $j$, and lightpaths diverging from path $1 \to i \to j \to 4$ at node $j$. $N_{th}$ is the total number of lightpaths that pass through link $L_{ij}$, $N_{tm}$ is the number of lightpaths terminating at node $j$, and $\frac{\Delta-2}{\Delta-1}(N_{th} - N_{tm})$ is the number of diverging lightpaths at node $j$. $\Delta = 6$ in the figure.

node to each other node, and each node employs the same routing policy. Then,

$$N_{th} = \frac{N_V(N_V - 1)\bar{H}}{2N_E}$$

$$N_{tm} = \frac{N_V - 1}{\Delta}$$

where $\bar{H}$ is the average number of hops over all paths. Because on average $P_{\lambda l}$ equals to the ratio of the number of lightpaths terminating or being switched to other paths

50

at node $j$ over the total number of lightpaths that pass through $L_{ij}$,

$$P_{\lambda l} = \frac{N_{tm} + \frac{\Delta-2}{\Delta-1}(N_{th} - N_{tm})}{N_{th}} \tag{3.2}$$

$$= \frac{2N_E}{\Delta(\Delta-1)N_V\bar{H}} + \frac{\Delta-2}{\Delta-1} \tag{3.3}$$

Denote $\rho_\lambda$ as the loading of wavelength channels. Then, Eq. (3.1) becomes

$$\rho_\lambda = \frac{P_{\lambda n}}{P_{\lambda n} + P_{\lambda l} - P_{\lambda n}P_{\lambda l}} \tag{3.4}$$

Solving the above equation for $P_{\lambda n}$, we obtain

$$P_{\lambda n} = \frac{\rho P_{\lambda l}}{1 - \rho(1 - P_{\lambda l})}$$

Assume shortest path routing. Define $P_{\lambda B}$ as the probability that a request cannot be served by network resources along the designated path at the time of the request. Define $P_{sB}$ as the conditional probability that a request, which cannot be served by network resources at the time of the request, also cannot be scheduled into the future as all the schedule holders on at least one link are full (Fig. 3-5).
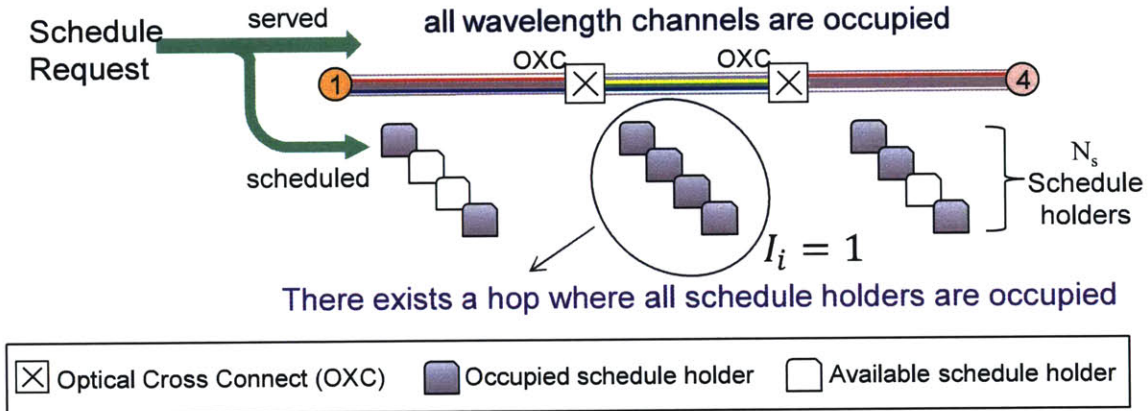


Figure 3-5: Illustration of the scenario where a request can be blocked.

Assume independence among lightpaths along the same path[5]. Because of wavelength continuity constraint, $P_{\lambda B}$ equals the probability that for any lightpath along

---

[5]This is a good assumption when there are many wavelength channels for low to medium loading, which will be the case as predicted in [26].

the designated path, there are interfering sessions. Therefore,

$$P_{\lambda B} = \Pr\{\text{All } \Lambda \text{ wavelength along the path are not available }\}$$
$$= (\Pr\{\text{A given wavelength along the path is not available}\})^{\Lambda}$$
$$= (1 - \Pr\{\text{A given wavelength is available on all hops along the path}\})^{\Lambda}$$
$$= \left[1 - (1 - P_{\lambda n})^{\bar{H}}\right]^{\Lambda}$$

where $\Lambda$ is the total number of wavelength on a link. This is a fair approximation of the expected $P_{\lambda_B}$ w.r.t distributions of $H$. The probability that a session enters a schedule holder is also the probability that a request is scheduled into the future and handled by a particular schedule holder, and,

$$P_{sn} = P_{\lambda n} P_{\lambda B} \frac{\Lambda}{N_s} \tag{3.5}$$

Because wavelength continuity is also respected by scheduled sessions, the probability that an existing session leaves the schedule holder of the corresponding wavelength channel is the same as the probability that the scheduled session leaves that lightpath, that is, $P_{sl} = P_{\lambda l}$. Define $\rho_s$ as the loading of schedule holders, from Eq. (3.1), we obtain

$$\rho_s = \frac{P_{sn}}{(P_{sn} + P_{sl} - P_{sn}P_{sl})} \tag{3.6}$$

Assume independence among schedule holders on the same link[6], and Markovian independence among neighboring hops[7] (the state of a designated link, given the state of the first neighboring link before it, is independent of states of all other links before its first neighbor). Define $P_{OB}$ as the overall probability that a request is blocked. Then,

$$P_{OB} = P_{\lambda B} P_{sB}$$

To obtain $P_{sB}$, we notice that a request that cannot be served by network resources at the time of the request is only blocked if, along the designated path, there exists a hop where all the schedule holders are occupied. Define an indicator random variable $I_i$ that equals to one if all schedule holders on hop $h_i$ are occupied, and zero otherwise.

---

[6]This assumption is good when loading of schedule holders is low, which is the case with large network demands, see Section 3.4.

[7]Same assumption as in [7].

Then,

$$P_{sB} = 1 - P\{I_1 = 0\} \prod_{i=2}^{\bar{H}} P\{I_i = 0 | I_{i-1} = 0\} \tag{3.7}$$

Assume a request encounters the steady state at schedule holders of the first hop. Then,

$$P\{I_1 = 0\} = 1 - \rho_s^{N_s} \tag{3.8}$$

From

$$P\{I_i = 1\} = P\{I_i = 1 | I_{i-1} = 0\}P\{I_{i-1} = 0\} + P\{I_i = 1 | I_{i-1} = 1\}P\{I_{i-1} = 1\}$$

we obtain,

$$
\begin{aligned}
P\{I_i = 0 | I_{i-1} = 0\} &= 1 - P\{I_i = 1 | I_{i-1} = 0\} \\
&= 1 - \frac{P\{I_i = 1\} - P\{I_i = 1 | I_{i-1} = 1\}P\{I_{i-1} = 1\}}{1 - P\{I_{i-1} = 1\}} \\
&= 1 - \frac{\rho_{si}^{N_s} - \rho_{s(i-1)}^{N_s}(1 - P_{sl} + P_{sl}P_{sn})^{N_s}}{1 - \rho_{s(i-1)}^{N_s}}
\end{aligned} \tag{3.9}
$$

where $\rho_{si}$ is the loading of schedule holders on hop $h_i$. Substituting Eq. (3.8) and (3.9) into Eq. (3.7) we get,

$$P_{sB} = 1 - (1 - \rho_s^{N_s}) \prod_{i=2}^{\bar{H}} \left[ 1 - \frac{\rho_{si}^{N_s} - \rho_{s(i-1)}^{N_s}(1 - P_{sl} + P_{sl}P_{sn})^{N_s}}{1 - \rho_{s(i-1)}^{N_s}} \right] \tag{3.10}$$

Because a request encounters the steady state loading of schedule holders at hop $h_1$, we will show in the following theorem that it also encounters the steady states of schedule holders at the following hops along the designated path.

**Theorem 1.** *For a $\bar{H}$-hop path in a symmetric network with uniform all-to-all traffic, if the loading of the schedule holders on $h_1$ is $\rho_s$, then the loadings of schedule holders of $h_2$, $h_3$,...,$h_{\bar{H}}$ are also $\rho_s$.*

53

Theorem 1 is proved in Appx. A.1, and Eq. (3.10) is reduced to

$$P_{sB} = 1 - (1 - \rho_s^{N_s}) \prod_{i=2}^{H} \left[ 1 - \frac{\rho_s^{N_s} - \rho_s^{N_s}(1 - P_{sl} + P_{sl}P_{sn})^{N_s}}{1 - \rho_s^{N_s}} \right]$$

$$= 1 - (1 - \rho_s^{N_s}) \left\{ 1 - \frac{\rho_s^{N_s} \left[ 1 - (1 - P_{sl} + P_{sl}P_{sn})^{N_s} \right]}{(1 - \rho_s^{N_s})^{\bar{H}-1}} \right\}^{\bar{H}-1} \quad (3.11)$$

where

$$P_{sl} = P_{\lambda l}$$

$$P_{sn} = P_{\lambda n} P_{\lambda B} \frac{\Lambda}{N_s}$$

Then, the overall blocking probability $P_{OB}$ can be written as:

$$P_{OB} = P_{\lambda B} \cdot P_{sB}. \quad (3.12)$$

Since the above derivation is based on a meshed network, to distinguish with the tunneled case, we denote the overall blocking probability of a network of Architecture M as, $P_{OBM}$[8], then

$$P_{OBM} = P_{\lambda B} \cdot P_{sB} \quad (3.13)$$

Figure 3-6 plots $P_{OBM}$ as a function of network loading $\rho$ with $N_s/\Lambda = 1$.

## 3.3  Blocking Probability for Architecture T

For Architecture T, since connections between any two nodes are tunneled, the hop count is one for all WAN paths, and schedule holders locates at the Ingress node of

---

[8]Subscript $M$ or $T$ is used to identify Architecture M or T.

54
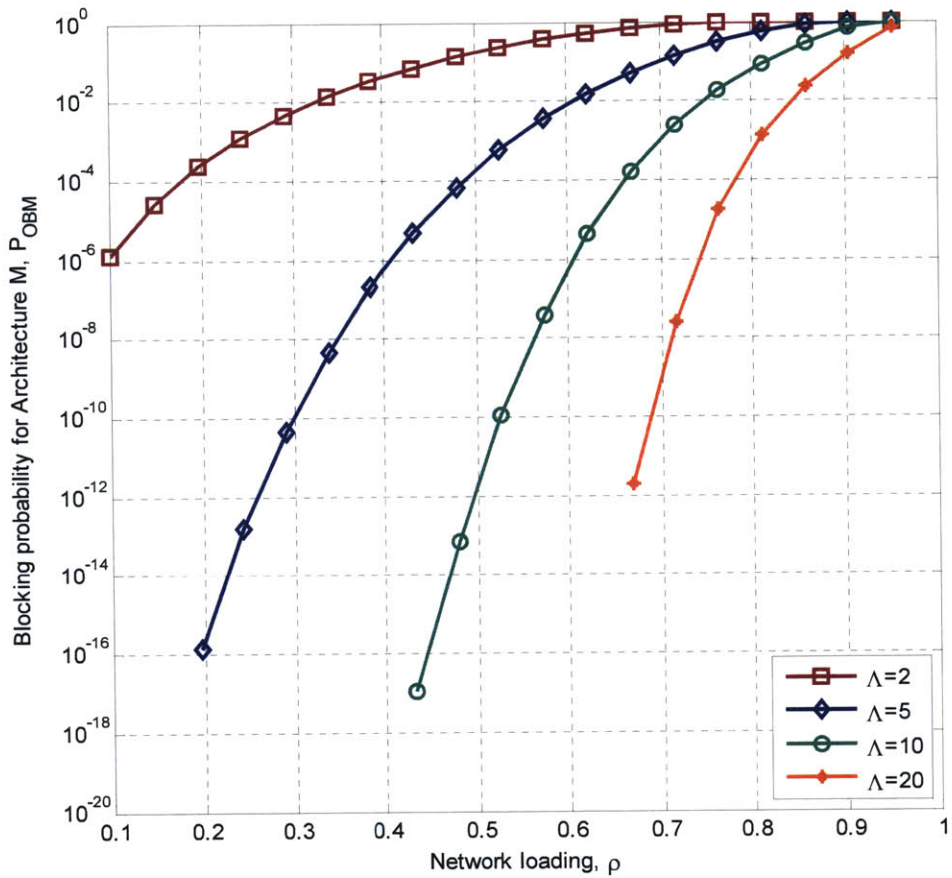
Figure 3-6: Blocking probability for Architecture M, $P_{OBM}$, w.r.t. network loading $\rho$ with $N_s/\Lambda = 1$, for different values of $\Lambda$. $N_s$ is the total number of schedule holders per link, and $\Lambda$ is the number of wavelengths per link.

WAN[9]. Therefore, we have

$$H = 1$$
$$N_E = N_V(N_V - 1)$$
$$\Delta = N_V - 1$$

---

[9]In the case of the tunneled architecture, since there is only one hop, all requests scheduled into the future in the same wavelength tunnel can be aligned back to back. Therefore, the queuing delay of a request can be approximated by the product of the average number of occupied schedule-holders-per-wavelength times the average transaction-time.

Therefore,

$$P_{\lambda l} = 1$$

$$P_{\lambda n} = \rho$$

$$P_{\lambda B} = \rho^{\Lambda_T}$$

$$P_{sl} = 1$$

Define $D$ as the average number of source-destination sessions over one link in the physical network topology. Then,

$$D = \frac{N_V(N_V - 1)\bar{H}}{2N_E} \tag{3.14}$$

Assume wavelength channels over one link in the physical network topology are shared evenly among all source-destination sessions that pass through that link using shortest path routing, then,

$$N_{sT} = \frac{N_s}{D}$$

$$\Lambda_T = \frac{\Lambda}{D}$$

where $N_{sT}$ is the number of schedule holders per tunnel, and $\Lambda_T$ is the number of wavelengths per tunnel.

Substituting the above equations into Eq. (3.5), (3.6), and (3.11), we obtain

$$P_{sn} = \frac{\Lambda_T}{N_{sT}}\rho^{\Lambda_T+1}$$

$$\rho_s = \frac{\Lambda_T}{N_{sT}}\rho^{\Lambda_T+1} \tag{3.15}$$

$$P_{sB} = \left(\frac{\Lambda_T}{N_{sT}}\right)^{N_{sT}} \rho^{N_{sT}(\Lambda_T+1)}$$

Denote $P_{OBT}$ as the overall blocking probability of Architecture T, then,

$$P_{OBT} = \left(\frac{\Lambda_T}{N_{sT}}\right)^{N_{sT}} \rho^{N_{sT}(\Lambda_T+1)+\Lambda_T} \tag{3.16}$$

Figure 3-6 plots $P_{OBT}$ as a function of network loading $\rho$ with $N_s/\Lambda = 2$.
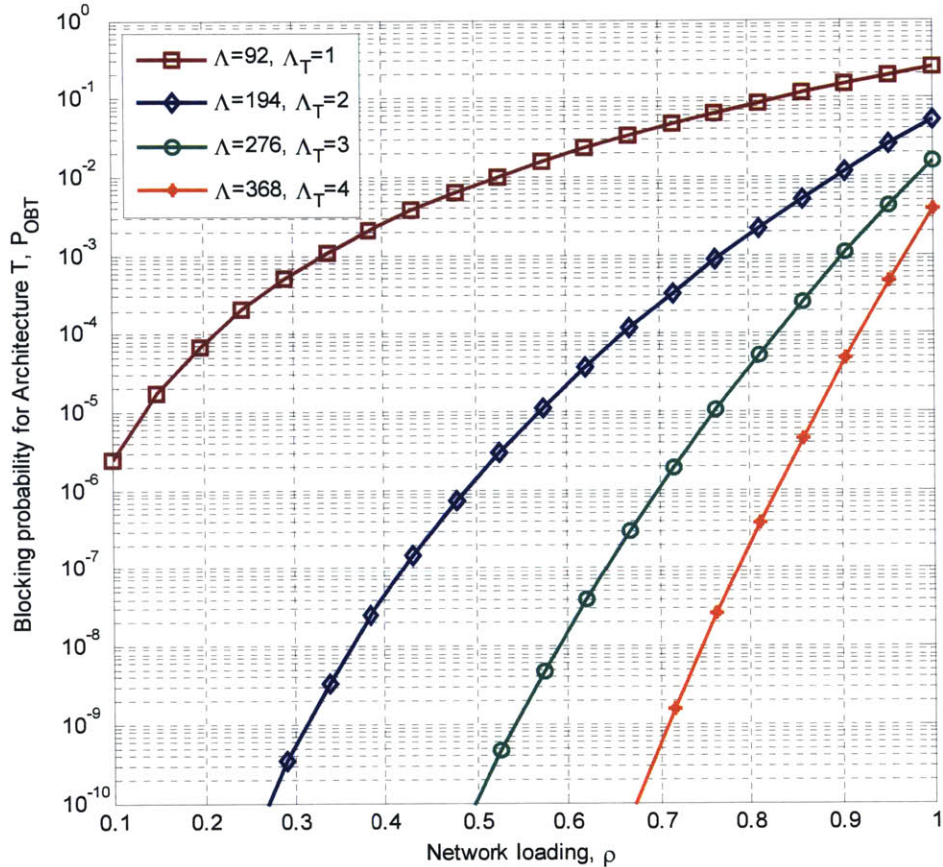
Figure 3-7: Blocking probability for Architecture M, $P_{OBT}$, w.r.t. network loading $\rho$ with $N_s/\Lambda = 2$, for different values of $\Lambda_T$. $N_s$ is the total number of schedule holders per link, $\Lambda$ is the number of wavelengths per link, and $\Lambda_T = \Lambda/92$ is the number of wavelengths per tunnel for the case where there are 92 tunnels per edge.

## 3.4 Performance Comparison of Architecture M and T

To compare data network performance of Architecture M and T, we consider the US backbone network and analyze the effects of network resources and schedule holders. Figure 3-8 [34, Fig. 8.1] shows the physical topology of the US backbone network and Table 3.1 summarizes its parameters. The US backbone network is a mesh network composed of 60 nodes. Each link in the figure represents about 100 fiber links, and for each fiber link there are about 200 wavelength channels. The average number of hops of an end-to-end connection is four. The average node degree is 2.6, while the largest node degree is five and the least node degree is two. Therefore, a typical path in the

backbone network comprises four links on the average, and there are interconnections at the connecting nodes, which introduce traffic merging and diverging.
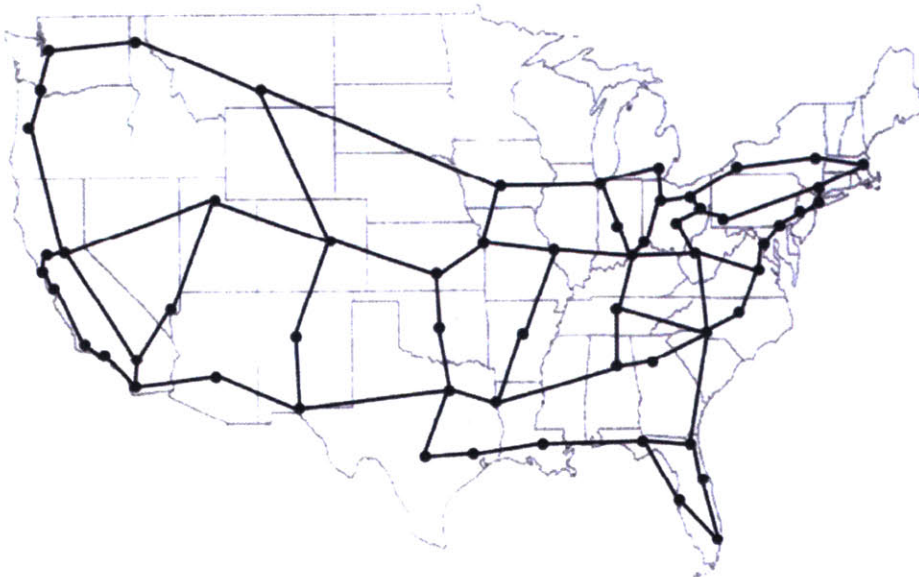


Figure 3-8: Optical backbone network of the United States. Reproduced from [34, Fig. 8.1].

For Architecture M, we have

$$N_V = 60$$
$$N_E = 77$$
$$\bar{\Delta} = 2.6$$
$$\bar{H} = 4$$

For Architecture T, since

$$D = \frac{N_V(N_V - 1)\bar{H}}{2N_E} \approx 92$$

Wavelength channels on each link of the physical topology are assigned equally to tunnels of 92 source-destination pairs. Therefore,

$$\Lambda_T = \frac{\Lambda}{92}$$

Using Eq. (3.13) and (3.16), and the above numerical values, we plot the blocking

58

Table 3.1: Important parameters for the US backbone network and their values, adopted from [34, Tbls. 8.1 and 8.2].

| Parameter | Value |
|---|---|
| Number of nodes | 60 |
| Number of links | 77 |
| Average node degree | 2.6 |
| Largest node degree | 5 |
| Least node degree | 2 |
| Average link length | 450 km |
| Number of wavelength channels per fiber link | 200 |
| Average number of hops of an end-to-end connection | 4 |

probabilities, throughput w.r.t. traffic demand, scheduling capacity and network loading. Fig. 3-9 plots blocking probabilities w.r.t. number of wavelength channels, for a fixed scheduling capacity of $N_s/\Lambda = 1$, and three different network loading conditions ($\rho = 0.6$, $\rho = 0.7$, and $\rho = 0.8$), for both the Meshed and Tunneled Architectures. For each fixed network loading (0.6, 0.7 or 0.8), when traffic demand increases, network resources (wavelength channels) also increase correspondingly. Fig. 3-9 shows that, even with medium to high loading, blocking probabilities for both architecture decreases quickly when traffic demand increases. When there are about three wavelength channels, with $N_s/V = 1$, both architecture can achieve a blocking probability that is less than $10^{-4}$.

Fig. 3-10, for a fixed network loading $\rho = 0.7$, plots the blocking probabilities for both architectures w.r.t. scheduling capacity for various traffic demands. It shows that network blocking probability decreases with more schedule capacity for both architectures. In particular, with high traffic demand ($\Lambda_T = 16$), blocking probability sharply decreases even when the $N_s/\Lambda$ is as small as 0.1. In fact, when $N_s/\Lambda \approx 1$, $\log P_{OBT}$ decreases approximately linearly w.r.t. $N_{sT}\Lambda_T$. Consider point A (in red font color) in

Fig. 3-10. At point A, we have

$$\Lambda = 194$$
$$\Lambda_T = 2$$
$$N_s/\Lambda = 2$$
$$\rho = 0.7$$

With the above values, we obtain $N_{sT}\rho_s < 1$. Therefore, with 0.7 network loading, and two wavelength channel of network resources per source-destination pair, the tunneled architecture achieves a blocking probability less than $10^{-3}$ and a network delay less than one transaction time[10].

For a fixed blocking probability of $10^{-4}$, Fig. 3-11 plots throughput w.r.t. the scheduling capacity $N_s/\Lambda$ for three different $\Lambda$ values (network resources). Throughput of Architecture T can be quickly increased when allowing more schedule holders in the system, and, when traffic demand is high (e.g. $\Lambda = 1000$ or $3000$), both architectures can achieve high loading with little scheduling capacity ($< 1$). Therefore, when traffic demand is high, with medium to high loading, the blocking probabilities of both architectures with the same amount of network resources, are much less than $10^{-4}$, which is a good target blocking probability of user agreement. And the performance of Architecture T can be greatly improved by allowing more schedule holders in the system.

---

[10]The amount of time a session spends in the schedule holder is proportional to the average delay by Little's Theorem. For most applications it is reasonable to have a delay of one transmission time and therefore the acceptable value of $N_s/\Lambda$ is from 1 to 4. When $N_s/\Lambda = 1$, the average delay is less than one transaction time.
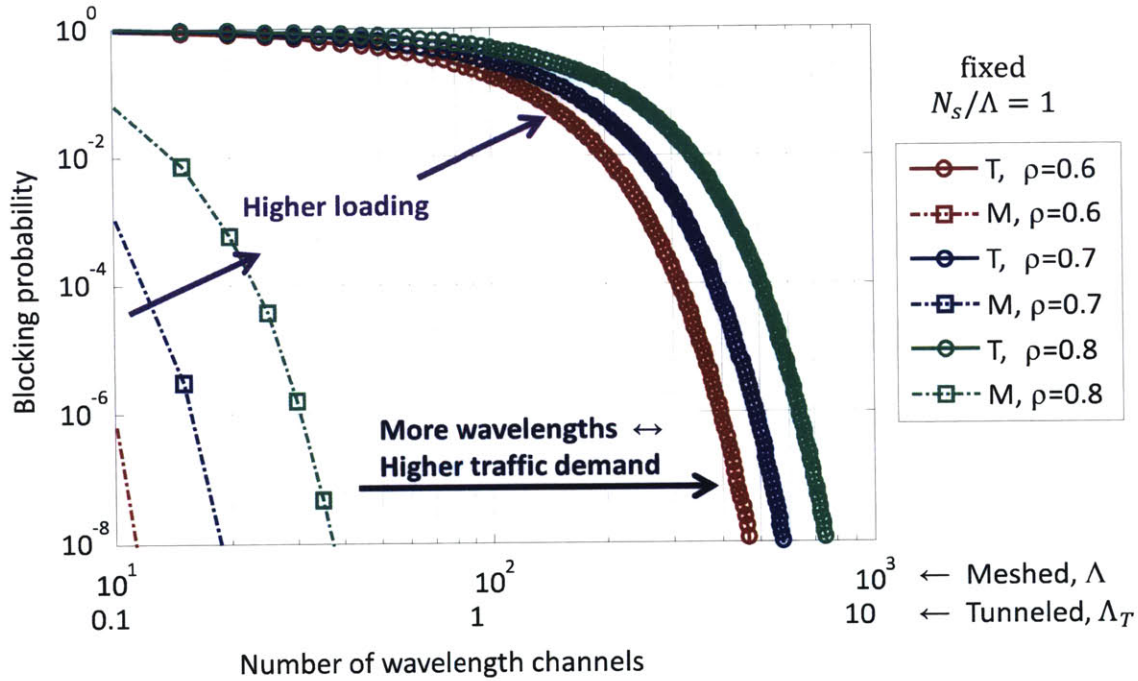
Figure 3-9: Blocking probability w.r.t. number of wavelength channels with $N_s/\Lambda = 1$, where $N_s$ is the total number of schedule holders per link, $\Lambda$ is the total number wavelengths per link in a meshed network topology, "M" is for Architecture M, and 'T" is for Architecture T. Because in the tunneled logical network topology, wavelengths on one link are divided equally to all $D$ source-destination tunnels going through that link, the average number of wavelengths for each source-destination tunnel is $\Lambda_T = \Lambda/D$. Therefore, two x-axis are used, with one for the number of wavelengths per link in Architecture M, and the other for the number of wavelengths per tunnel for Architecture T.
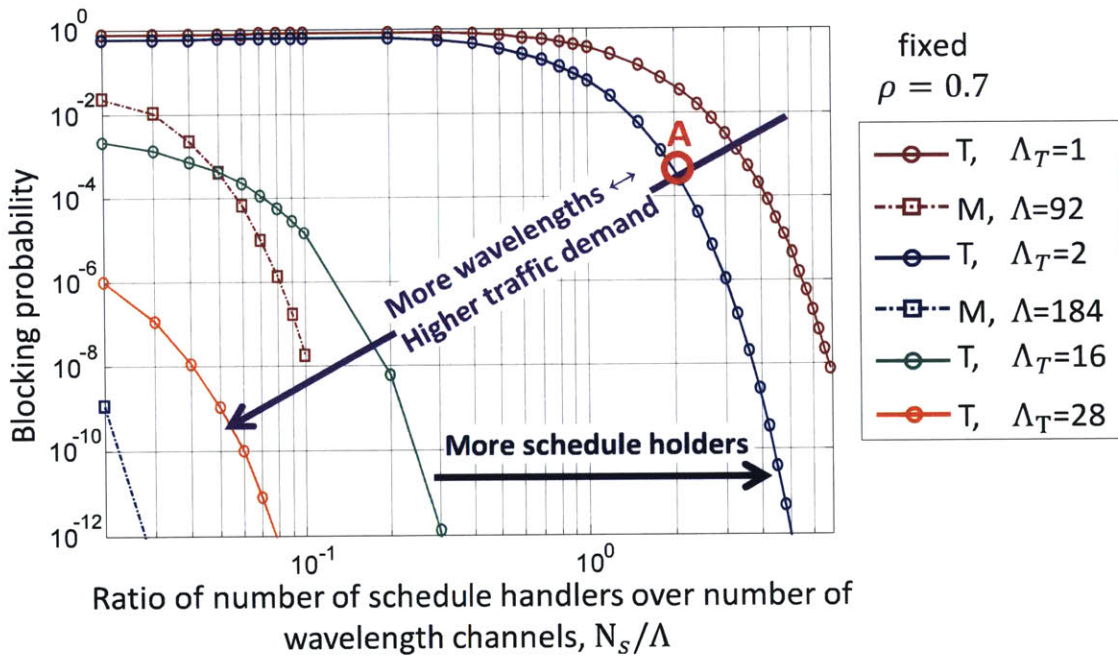
Figure 3-10: Blocking probability w.r.t. ratio of number of schedule holders and wavelength channels with network loading $\rho = 0.7$ for cases where the number of wavelengths per tunnel for Architecture T $\Lambda_T = 1, 2, 16$ and 28, and the number of wavelengths per link for Architecture M $\Lambda = 92, 184, 1472$ and 2600, respectively. The blocking probability for the meshed case with $\Lambda = 1472$ and $\Lambda = 2600$ are below $10^{-12}$ and are not shown in the figure. The meshed counterpart of the orange solid curve (T, $\Lambda_T = 28$) corresponds to $\Lambda = 2600$, the projected traffic demand per WAN node in [26].
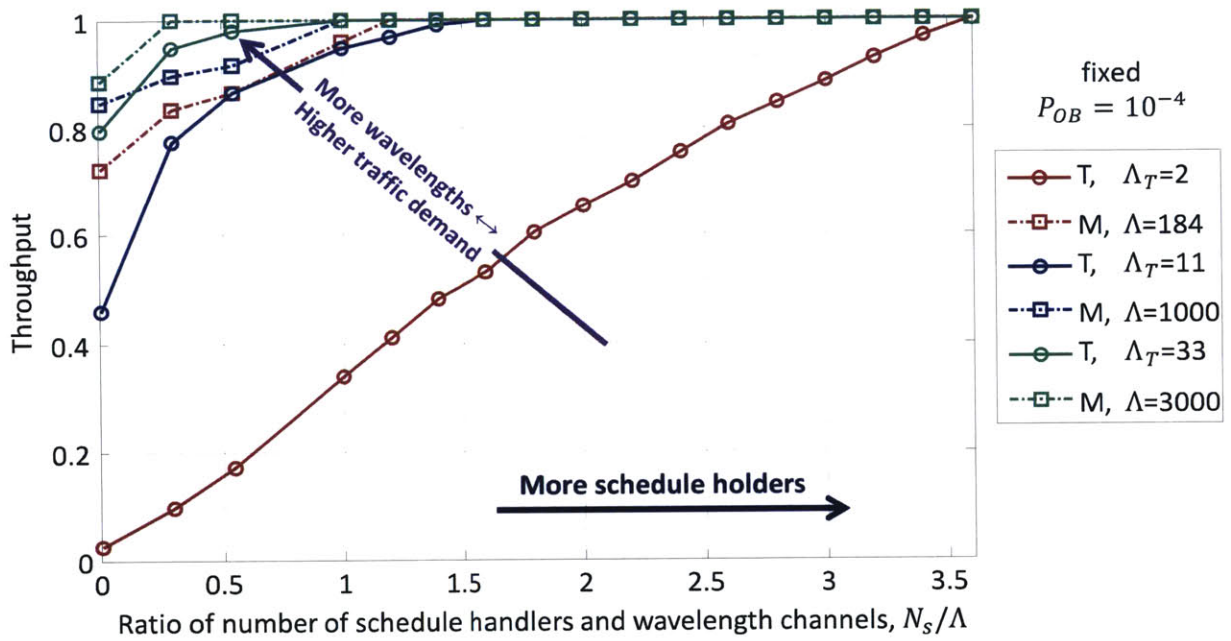
Figure 3-11: Throughput w.r.t. ratio of number of schedule holders and wavelengths with $P_{OB} = 10^{-4}$ for cases where the number of wavelengths per tunnel for Architecture T $\Lambda_T = 2, 11$, and 33, and the number of wavelengths per link for Architecture M $\Lambda = 184, 1000$, and 3000, respectively. Note that not all number of wavelengths per link can be divided to be assigned to tunnels, the number of wavelengths in the figure were chosen in a way such that real number of throughput can be obtained.

# Chapter 4

# The Scheduling Algorithm

When a request arrives at a Scheduler, the Scheduler runs the scheduling algorithm to decide when, along which path, in which wavelength channel the request is going to be transmitted. We assume there are no wavelength converters along a transmission path and the same wavelength is used on all hops along the path. The scheduling therefore is done in a way such that there is no collision of transmission using the same wavelength channel in the network. In this chapter, we briefly discuss the related literature in Section 4.1. We introduce scheduling of optical flow switching in Section 4.2 and 4.3, discuss the optimal scheduled wavelength assignment problem in Section 4.4, and propose a heuristic algorithm for scheduled wavelength assignment in Section 4.6. We then analyze the control traffic in Section 4.7.

## 4.1 Routing and Wavelength Assignment in All-Optical Networks

The Routing and Wavelength Assignment (RWA) problem in all-optical networks arose when emerging applications requiring higher bandwidth called for a more efficient use of the network resources in the 1990's [13], and RWA is a network optimization formulation to provide lightpath connection for traffic demand while minimizing network cost for a given set of constraints. Given a network topology and a set of traffic demands (in the unit of lightpath), the RWA problem determines a path and a wavelength channel for each traffic demand in the network. If there is no wavelength converters, there is a wavelength continuity constraint for lightpath set-ups, that is, along a path from the source node to the destination node, the same wavelength needs to be assigned to the

lightpath on all the links. If for a certain request, no lightpath can be set up along candidate paths between the source and the destination nodes, the request is said to be blocked. Numerous studies formulated the RWA problem into Integer Linear Programming (ILP) problems with variants of objectives, such as, to minimize the number of wavelength channels used to meet a target blocking probability, or to minimize the blocking probability with limited network resources [31, 44, etc.]. However, the ILP problems are generally NP-hard [18] and take too much time to solve. Therefore, the RWA problem in many studies was decomposed to into a lightpath routing problem and a wavelength assignment problem. The former problem of lightpath routing, depending on the objectives, had often been formulated into network flow problems. Network flow problems with integer constraints are again NP-hard, and various heuristic algorithms have been developed, see for example, [32, 24]. The wavelength assignment problem, when being formulated to either minimize the number of wavelength channels or to minimize the blocking probabilities, is NP-complete [14]. Therefore, for medium to large networks, heuristics have been studied [6, 44]. A large portion of the literature on RWA studied the problem for the network planning phase, that is, the lightpath demands are known beforehand, and the computation usually can be done off-line during network planning phase. This type of RWA problem was named the static RWA problem. In a situation when lightpath requests arrive randomly, the RWA problem computes lightpath assignment on-the-fly, and if there is no network resource available at the time of the request, the request is blocked. This later case is often named as the dynamic lightpath establishment problem [46]. For readers who are interested in learning more about the literature of RWA, [46] and [17] both provided comprehensive literature reviews on this topic.

Most of the literature dealt with lightpath demands with known transmission time (the lightpath request comes with information when the lightpath needs to be set up and when it can be torn down). Gagnaire et al. in [17] studied the lightpath establishment problem for mixed traffic of three types: Permanent Lightpath Demands (PLD), Scheduled Lightpath Demands (SLD), and Random Lightpath Demands (RLD). The PLD refers to lightpath connections that are permanent and is computed off-line during network planning phase. The SLD refers to connection requests with known set-up and tear-down times. Heuristics algorithms have been studied to either compute the SLD off-line in a global optimization problem formulation, or to compute the SLD sequentially ranked by the start-time of each SLD. RLD refers to requests that comes

up randomly and the arrival time is not known beforehand, but once an RLD traffic arrives at the network, the duration is known. Therefore, once an RLD arrival occurs, the RWA problem needs to compute a lightpath assignment on-the-fly that can accommodate this request from the time of arrival to the declared life time. Note that although, Gagnaire et al. named the lightpath demand as scheduled or random lightpath demand, the problem they formulated are different from the scheduling problem in optical flow switching. Because in their formulation, the traffic transmission time is declared in the request, and therefore is fixed[1]. In our scheduling problem, in terms of transmission time, a user request only declare the file size (or, equivalently, the transmission duration), and it is up to the Scheduler to assign a lightpath and ALSO a scheduled transmission time (start time and end time) to each request. Therefore, the variants of RWA problems in the literature are different from what we are studying here for optical flow switching.

## 4.2 The Scheduling Problem for Optical Flow Switching

In [40], Weichenberg et al. first studied the scheduling algorithm for optical flow switching where a user request can tolerate a certain amount of delay and a request is scheduled to sometime in the future if it can not be served at the time of the request. Recognizing the complexity of an optimal scheduling algorithm, Weichenberg et al. proposed a heuristic scheduling algorithm over a network topology with tunneled WAN and broadcast MAN and Access Network. Requests are queued at the ingress WAN node and processed on a first-in-first-out basis in two stages, that is, WAN resources were considered first, and then MAN and Access Network. Due to the higher cost of WAN resources, resources in the WAN were designed to be used more efficiently (with high loading) and their availability dictates which wavelength channel can be selected. The scheduling algorithm then check and wait for the availability of the same wavelength channel in the source and destination MANs and Access Networks, and reserve the channel once it becomes available for the request.

In this thesis, we focus on OFS scheduling in the wide area network and investigate the algorithm complexity and control traffic efforts with respect to not only traffic

---

[1]This tends to cause high blocking probability at all loading levels due to inflexibility of the start time.

demand, network resources, but also network topology. In particular, we will analyze the impact of two example topologies, the Meshed topology and the Tunneled topology, on scheduling complexity and control traffic efforts.

## 4.3   The Shortest Path Routing Algorithm

As in the literature where the RWA problem is decomposed into the routing problem and the wavelength assignment problem to make the problem tractable, in our heuristic algorithms, we also decompose the OFS scheduling problem into the routing problem and the scheduled wavelength assignment problem.

For the routing problem, we use shortest-path routing (SPR), where the shortest path is the path with the least number of hops[2]. We choose shortest path routing algorithm for its simplicity, and we argue that with adaptive-to-traffic-demand network resources deployment or lighting-up new fibers, shortest-path routing is also energy and cost-efficient. For the short term when traffic increases unpredictably quickly and if new network resources cannot be deployed immediately, the second shortest-path can be used to route the incremental traffic. Note that with careful planning and traffic monitoring, this unpredicted rapid traffic increase should be a rare situation. Because the WAN topology does not change frequently, the SPR algorithm only runs when there are changes to the logical network topology and all pair shortest paths are stored either centrally when the algorithm is centralized or, otherwise, at each node. We use the Floyd-Warshall algorithm [16, Chap.25.2] to calculate all-pair shortest paths for a graph $G(V, E)$ with running time $\Theta(N_V^3)$, where $N_V$ is the number of nodes in $V$.

## 4.4   The Scheduled Wavelength Assignment Problem

With lightpath routing determined by the Floyd-Warshall Algorithm in Section 4.3, what remains to be solved is the scheduled wavelength assignment problem (SWA). The scheduling wavelength assignment problem for optical flow switching must satisfy the following constraints:

---

[2]We choose the shortest path to be the path of the least number of hops because this will result in minimum amount of control traffic for lightpath set-ups and tear-downs as the number of switches that need to be configured is minimized.

1. Two requests with overlapping transmission time must not be assigned the same wavelength on a given link.

2. With no wavelength conversation, the same wavelength must be assigned to a lightpath on all links along its path.

The wavelength assignment problem in the literature, when being formulated to minimize the number of total wavelengths required, has been shown to be NP-complete [14]. We briefly outline the proof here. The essence of proof lies on the equivalence between

- The $\omega$-wavelength-assignment problem whether wavelength assignment for a set of routed lightpath over a network topology graph $G$ can be achieved using $\leq \omega$ wavelengths, and,

- The $n$-graph-colorability problem whether a graph (named it the path graph $G_P$) can be colored using $\leq n$ colors so that no two vertices sharing an edge are assigned the same color.

The proof includes two parts:

(a) The $\omega$-wavelength-assignment problem can be reduced, in polynomial time, into a $\omega$-graph-colorability problem, and therefore, solving the $\omega$-graph-colorability problem solves the $\omega$-wavelength-assignment problem. This is achieved by, for a given list of routed lightpath $\mathbf{L}$ over a network topology $G$, creating a path graph $G_P(V_P, E_P)$ in a way as:

  - For each lightpath $l$ in $\mathbf{L}$, create a corresponding vertex $v_l$ in $V_P$.

  - For two vertices $v_i$ and $v_j$ in $V_P$, create an edge $e_{i-j}$ in $E_P$ if the corresponding lightpath $i$ and $j$ in $\mathbf{L}$ share some link over $G$.

(b) The $n$-graph-colorability problem can be reduced, in polynomial time, into an $n$-wavelength-assignment problem, and therefore, solving the $n$-wavelength-assignment problem solves the $n$-graph-colorability problem.

Therefore, after parts (a) and (b), the equivalence between the $\omega$-wavelength-assignment problem and the $n$-graph-colorability problem has been proven. Since the $n$-graph-colorability problem is NP-complete [18, 15], so is the $\omega$-wavelength-assignment problem.

69

Define the optimal scheduled wavelength assignment to be the assignment that minimizes the total number of wavelengths required for a set of requests that arrive during the interested time horizon $T_h$. In optical flow switching, since user requests arrive randomly in the form of $\langle v_s, v_d, t_d \rangle$ with different arrival times, $t_a$. Therefore, a request can be specified by the quadruple $\langle v_s, v_d, t_d, t_a \rangle$. Define $\mathbf{R}$ to be the list of all request arrivals during time $T_h$. In practice, any on-line scheduled wavelength assignment algorithm cannot be optimal for all possible instances of the stochastic arrivals. In theory we assume the optimal algorithm performs wavelength assignment with full knowledge of future traffic arrivals, i.e., the list of arrivals, $\mathbf{R}$. Define list $\mathbf{T_t} = \left[ \langle t_{s_1}, t_{e_1} \rangle, \langle t_{s_2}, t_{e_2} \rangle, \cdots, \langle t_{s_{|\mathbf{L}|}}, t_{e_{|\mathbf{L}|}} \rangle \right]$ to be the list of scheduled transmission times of the result of the optimal schedule, where $|\mathbf{L}|$ is the total number request arrivals during $T_h$. Notice that if the optimal $\mathbf{T_t}$ is known, then we can also establish an equivalence between the scheduled wavelength assignment problem and the graph coloring problem. The proof is similar to the proof above with two differences:

a) The scheduled $\omega$-wavelength-assignment problem with known $\mathbf{T_t}$ can be reduced, in polynomial time, into a $\omega$-graph-colorability problem, in a way as:

  - For each lightpath $l$ in $\mathbf{L}$, create a corresponding vertex $v_l$ in $V_P$.

  - For two vertices $v_i$ and $v_j$ in $V_P$, create an edge $e_{i-j}$ in $E_P$ if the corresponding lightpath $i$ and $j$ in $\mathbf{L}$ share some link over $G$ and do not overlap in transmission times.

b) The $n$-graph-colorability problem can be reduced, in polynomial time, into an scheduled $n$-WA problem. Notice that the WA problem is a special case of the scheduled WA problem where the transmission time of all lightpath overlaps.

Therefore, the scheduled $\omega$-wavelength-assignment problem with known $\mathbf{T_t}$ is NP-complete. Hence we have the theorem below.

**Theorem 2.** *The optimal scheduled wavelength assignment problem with known future traffic arrivals is at least NP-complete.*

### 4.4.1 Complexity of the Optimal Scheduled Wavelength Assignment Algorithm with Known Future Traffic Arrivals

Although the optimal Scheduled WA problem is NP-complete, we can approximate the magnitude of the complexity to solve this problem. Define $\Lambda$ to be the total number

of wavelengths in the network, then with $|\mathbf{L}|$ request arrivals during $T_h$, the total number of possible wavelength assignments is $\Lambda^{|\mathbf{L}|}$. Therefore, an optimal algorithm that computes the optimal over all possible wavelength assignments takes running time $O(\Lambda^{|\mathbf{L}|})$. The average processing power over $T_h$ is, therefore, at least $O(\Lambda^{|\mathbf{L}|})/T_h$.

For Architecture T, because tunneling results into decoupling of the problem, the complexity is $O\left(\left(\frac{\Lambda}{D}\right)^{\frac{|\mathbf{L}|}{N_V(N_V-1)}}\right)$, where $D$ is the average number of source-destination tunnels over one link as defined in Eq. (3.14). For the WAN topology in Fig. 3-8, $D \approx 92$, and $N_V = 60$. Assume $\Lambda = 200$, then $O\left(\left(\frac{\Lambda}{D}\right)^{\frac{|\mathbf{L}|}{N_V(N_V-1)}}\right) \approx O\left(2^{\frac{|\mathbf{L}|}{3540}}\right)$, which is solvable if $|L|$ is not too large. Therefore, tunneling greatly simplifies the optimal algorithm, by magnitude of $D^{\frac{|L|}{N_V(N_V-1)}}\left(\Lambda^{1-\frac{1}{N_V(N_V-1)}}\right)^{|L|}$.

# 4.5 The Mathematically-Optimal Algorithm for Scheduled Wavelength Assignment

When future traffic arrivals are stochastic and unknown, one common way (in both industry and academia) of formulating the scheduled wavelength assignment problem is, for each new request, to minimize the number of wavelengths assigned in the network by searching over all possible routing and wavelength assignments and choosing an optimal schedule assignment. Let's name an algorithm that find the mathematically-optimal schedule assignment with the aforementioned formulation the Mathematically-Optimal (MO) Algorithm. However, because with a meshed graph, even the problem of finding the longest simple path[3] is NP-hard [33], the problem of searching over all possible routing and wavelength assignment to find the mathematically-optimal schedule assignment is at least NP-hard.

## 4.5.1 Complexity of the Mathematically-Optimal Algorithm

With the Meshed Architecture which has a network logical topology graph $G(V, E)$, there are in total $2^{N_E}$ possible paths between a source and a destination node, where $N_E$ is the number of edges in $G(V, E)$. Therefore, the running time of the Mathematically-Optimal Algorithm with Architecture M is $O(2^{N_E} F_M \Lambda_M)$, where $F_M$ and $\Lambda_M$ are the numbers of fibers and wavelengths per edge. However, with the Tunneled Architecture, the scheduling of traffic for one source-destination pair is decoupled with traffic for the

---

[3]A path is simple if it does not have any repeated vertices.

other source-destination pairs, the routing and wavelength-assignment problem degenerates and the Mathematically-Optimal Algorithm only needs to search over wavelength tunnels assigned to that specific source-destination pair. Therefore, the running time of the Mathematically-Optimal Algorithm with Architecture T is $O(F_T \Lambda_T)$ where $F_T$ and $\Lambda_T$ are the numbers of fibers and wavelengths per tunnel.

## 4.6 The FIFO-EA Algorithm for Scheduled Wavelength Assignment

The scheduled wavelength assignment problem selects a wavelength channel that is available on all the links along the shortest path during the whole period of the scheduled transmission time. For simplicity, we adopt the first-in-first-out (FIFO) principle, and selects the smallest number-ed wavelength channel with the earliest available time. Name this wavelength assignment algorithm the First In First Out Earliest Available (FIFO-EA) Algorithm, and the overall scheduling algorithm the SPR-FIFO-EA Algorithm.

---

**Algorithm 4.1** The FIFO-EA Algorithm

---

FIFO-EA$(R, P, t_d)$
1  **for** $\omega \leftarrow 1$ **to** $\Lambda_0$
2       $Sch(\omega) \leftarrow$ COLORPATH-EA$(P, R, t_d, \omega, t_0)$
       // Select the Schedule with the earliest available time
       // among those of all wavelengths.
3  $SCH \leftarrow \min(Sch.T_s)$
4  **if** $SCH.t_s = \infty$
5       **return** Null
6  **else return** $SCH$

---

$G(V, E)$ represents the logical topology of the WAN. Request $R$ is in the form of $\langle v_s, v_d, t_d \rangle$, where $v_s$ is the source node, $v_d$ is the destination node, and $t_d$ is the requested duration of the connection. $P$ is the shortest path from $v_s$ to $v_d$, and is composed of $H$ hops including $\langle h_P^1, h_P^2, ..., h_P^H \rangle$. Let $\Lambda_0$ be the number of wavelength channels along one path[4]. Let $f_{P\omega}^k$ be the number of fibers on link $l_P^k$ with wavelength $\omega$, and $F$ for the

---

[4] $\Lambda_{0M}$, the $\Lambda_0$ in Architecture M, is the same as $\Lambda$, the number of wavelengths per edge; while $\Lambda_{0T}$, the $\Lambda_0$ in Architecture T, is the same as $\Lambda_T$, the number of wavelengths per tunnel.

**Algorithm 4.2** The COLORPATH-EA Subroutine

COLORPATH-EA$(P, R, t_d, \omega, t_0)$

1  $Sch_1 \leftarrow$ LATESTMIN-OALINKS$(P, R, t_d, \omega, t_0)$
2  $Sch_2 \leftarrow$ LATESTMIN-OALINKS$(P, R, t_d, \omega, Sch_1.T_s)$
3  **if** $Sch_1.T_s = Sch_2.T_s$
4      **return** $Sch_2$
5  **else return** COLORPATH-EA$(P, R, t_d, \omega, Sch_2.T_s)$

---

**Algorithm 4.3** The LATESTMIN-OALINKS Subroutine

LATESTMIN-OALINKS$(P, R, t_d, \omega, t_0)$

1  $Link \leftarrow P.Link$
2  $v_d \leftarrow P.v_d$
3  **for** $h \leftarrow 1$ **to** $Link.Hops$
4      $lSch(h) \leftarrow$ MIN-OAFIBERS(
                    $P, R, t_d, Link(h).Fibers, t_0)$
5  $v_d Sch \leftarrow$ MIN-OATIMESEGMENTS$(v_d.Usage, t_d, t_0)$
6  $Sch.T_s \leftarrow \max(lSch.T_s, v_d Sch.T_s)$
7  $Sch.FiberIndex \leftarrow lSch.FiberIndex$
8  **return** $Sch$

---

**Algorithm 4.4** The MIN-OAFIBERS Subroutine

MIN-OAFIBERS$(P, R, t_d, Fibers, t_0)$

1  **for** $f \leftarrow 1$ **to** $Fibers.Size$
2      $fSch(f) \leftarrow$ MIN-OATIMESEGMENTS(
                    $Fibers(f).usage, t_d, t_0)$
3  $Sch.T_s \leftarrow \min(fSch.T_s)$
4  $Sch.FiberIndex \leftarrow \min(fSch.T_s).Index$
5  **return** $Sch$

**Algorithm 4.5** The MIN-OATIMESEGMENTS Subroutine

---

MIN-OATIMESEGMENTS($usage, t_d, t_0$)

1  **if** $usage.size < N_s$
2      $usgPnt \leftarrow usage.Headead$
       // Initialize $usgPnt$ to point to head of $usage$.
3      **while** $usgPnt$ NOT $usage.End$
4          $(t_s, t_e) \leftarrow usage(usgPnt).segment$
5          **if** $t_0 < t_s$
6              **if** $t_0 + t_d < t_s$
7                  $Sch.T_s \leftarrow t_0$
8                  **return** $Sch$
9          $usgPnt \leftarrow usage(usgPnt).next$
10         $t_0 \leftarrow t_e$
11     $Sch.T_s \leftarrow t_0$
12     $usage \leftarrow$ **Null**
13 **else** $Sch.T_s \leftarrow \infty$
14 **return** $Sch$

---

total number of fibers on a link. Time usage with $S$ segments ($S \leq N_s$) on a particular wavelength on a particular fiber on a link is in the form of $\langle (t_s^1, t_e^1), (t_s^2, t_e^2), ..., (t_s^S, t_e^S) \rangle$. Pseudo-code of the FIFO-EA Algorithm is given in Algorithm 4.1.
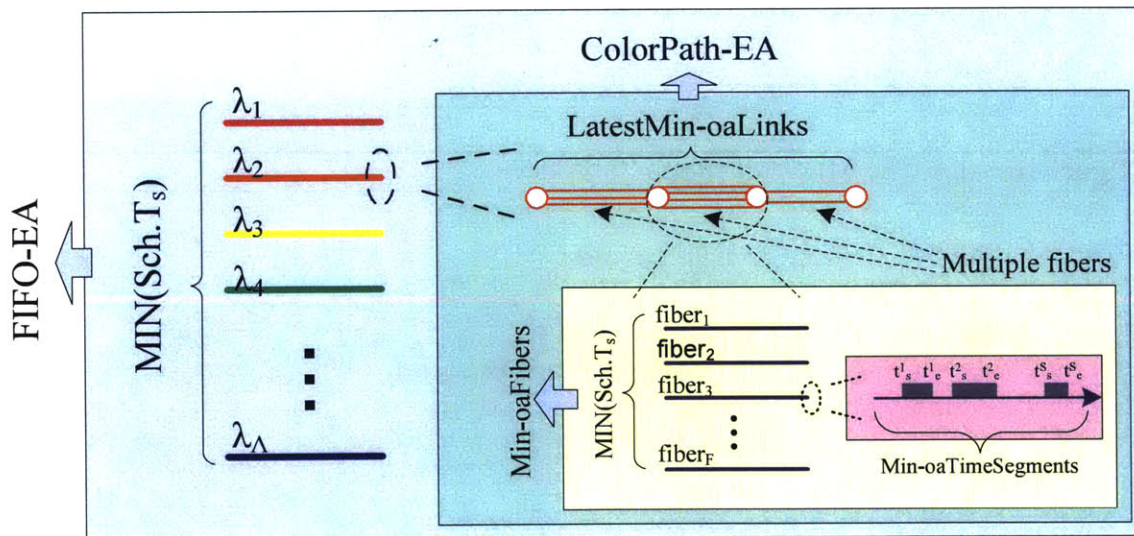


Figure 4-1: Illustration of Algorithm FIFO-EA.

74
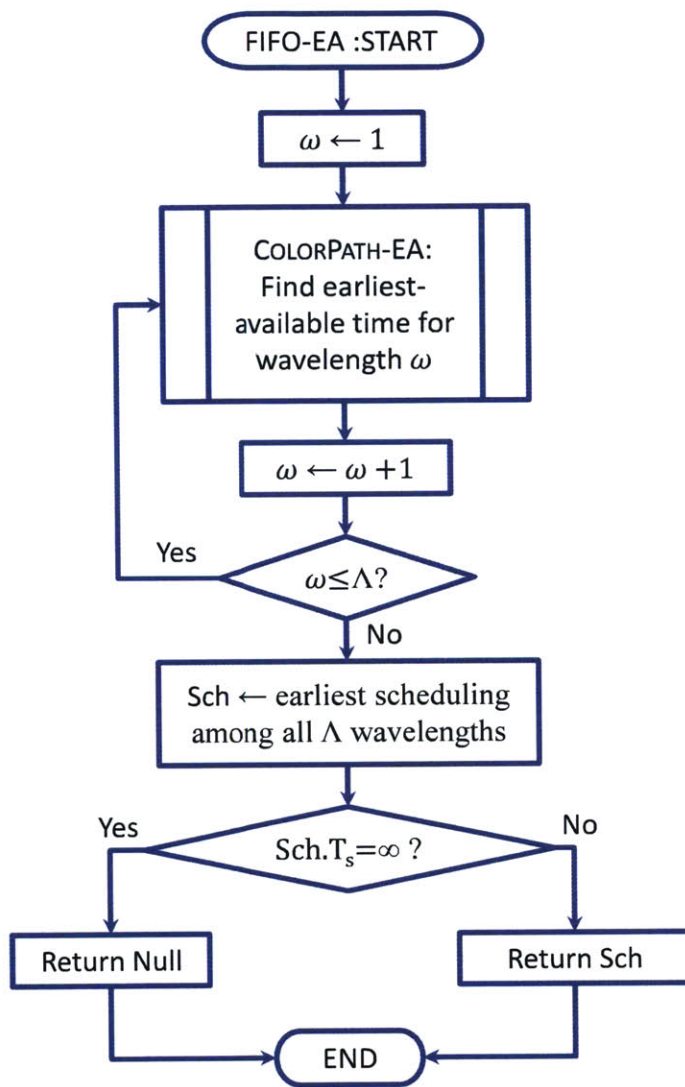
Figure 4-2: Flowchart for Algorithm FIFO-EA.
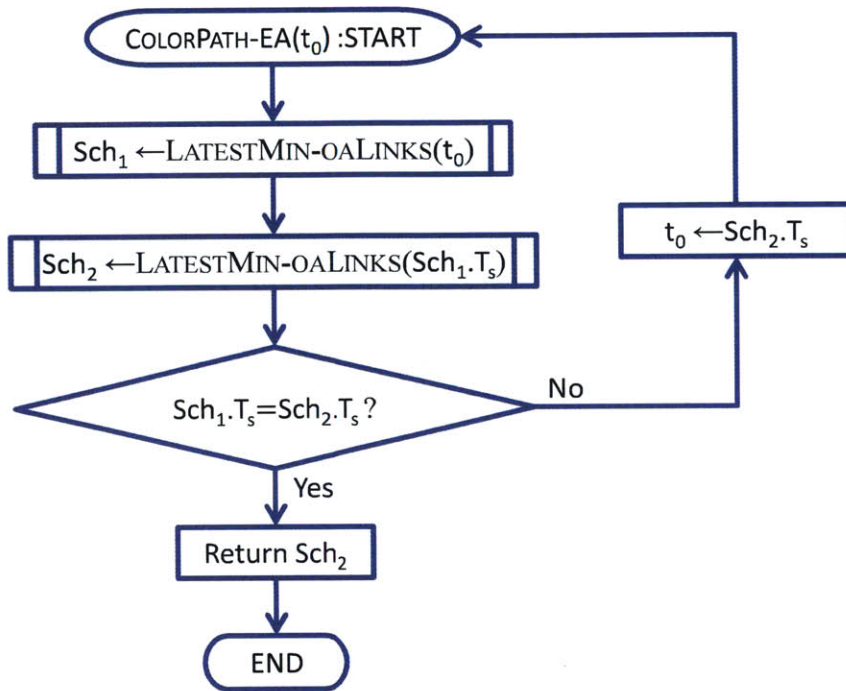
Figure 4-3: Flowchart for Subroutine COLORPATH-EA.
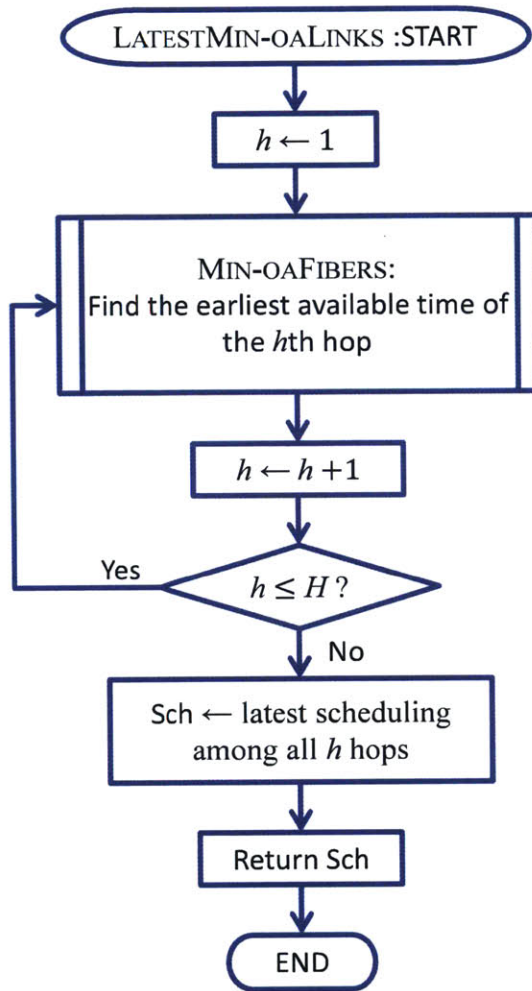
Figure 4-4: Flowchart for Subroutine LatestMin-oaLinks.
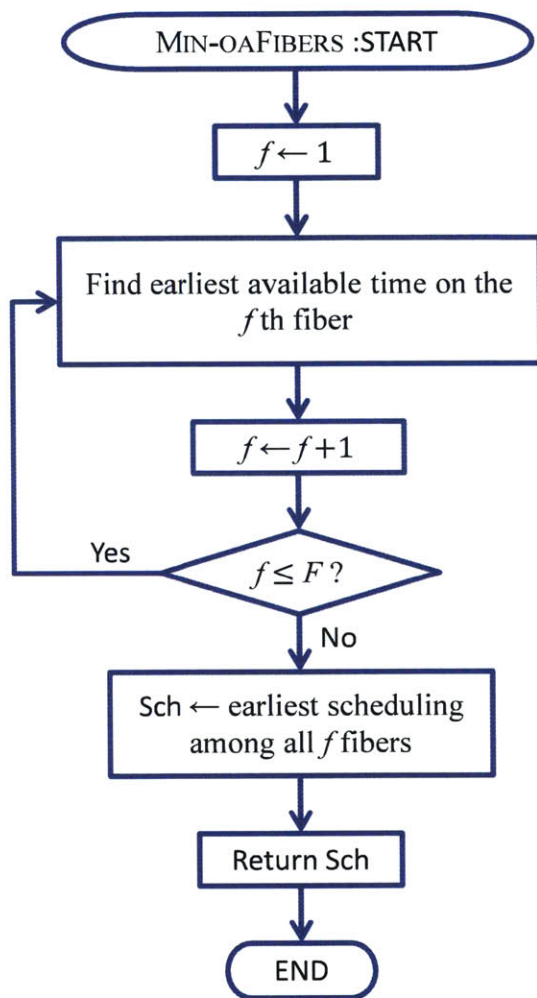
Figure 4-5: Flowchart for Subroutine MIN-OAFIBERS.

Fig. 4-1 and flowcharts in Fig. 4-2, 4.2, 4-4 and 4-5 illustrate the flow of the FIFO-EA Algorithm. The routine FIFO-EA takes input $(R, P, t_d)$, and iterates through all $\Lambda_0$ wavelengths and outputs, if the request can be accommodated, a schedule with the earliest available time among the individual earliest available time over each wavelength channel or, otherwise, "$\infty$" (the request is blocked). The earliest available time on one wavelength is determined by the subroutine COLORPATH-EA (Algorithm 4.2). The intuition behind COLORPATH-EA is that if the $i$th link of path $P$ is not available until time $\tau$ then path $P$ is not available until time $\tau$. In other words, the earliest available time of path $P$ cannot be earlier than the earliest available time of any link on the path. If link $i$ is not available until time $\tau$, to examine the available time of any other link, we only need to examine usages from time $\tau$ onwards. On the

other hand, if $\tau$ is the earliest available time of path $P$, $\tau$ is also the earliest available time of at least one link on the path, and the earliest available time of any links over path $P$ starting from $\tau$ is $\tau$. Therefore, in COLORPATH-EA, the latest of the earliest available time on each link over the path starting from the time at the request is first determined (let it be $Sch_1.T_s$); then the latest of the earliest available time on each link over the path starting from $Sch_1.T_s$ is determined (let it be $Sch_2.T_s$). If $Sch_1.T_s$ is the same as $Sch_2.T_s$, then all links are available starting from $Sch_1.T_s$ and it is the earliest available time of the path. Otherwise, the path cannot be available before $Sch_2.T_s$, and COLORPATH-EA is called recursively starting from $Sch_2.T_s$.

Subroutine LATESTMIN-OALINKS (Algorithm 4.3) first calculates the earliest available time over each individual link, the earliest available time at the destination node, and then outputs the latest of the earliest available times over all links and at the destination node. The earliest available time over an individual link is determined by subroutine MIN-OAFIBERS (Algorithm 4.4), where the earliest available time is determined for each fiber on the link for the corresponding wavelength, and the earliest among the earliest available times for each fiber is output. The earliest available time over one fiber on a link for a corresponding wavelength is determined by MIN-OATIMESEGMENTS (Algorithm 4.5), where the usage is checked to output the earliest available time when the channel on that fiber is available for the duration $t_d$.

## 4.6.1 Running Time Analysis of the FIFO-EA Algorithm

Let $T_1(S)$ be the running time of FIFO-EA, $T_2(S)$ be the running time of COLORPATH-EA, $T_3(S)$ be the running time of LATESTMIN-OALINKS, $T_4(S)$ be the running time of MIN-OAFIBERS, and $T_5(S)$ be the running time of MIN-OATIMESEGMENTS. The worst case happens when every time segment on the path is checked. Thus, for the worst case,

$$T_1(S) = \Lambda_0 T_2(S) + \Lambda_0 \tag{4.1}$$

$$T_2(S) \leq 2T_3(S) + T_2(S - \delta) \tag{4.2}$$

$$T_3(S) = HT_4(S) + T_5(S_{v_d}) + H + 1 \tag{4.3}$$

$$T_4(S) = FT_5(S) + F \tag{4.4}$$

where $\delta$ is the number of time segments that have already been checked. Substituting Eq. (4.3) and (4.4) into (4.2), we obtain $T_2(S) = O(SHF)$. Hence,

$$T_1(S) = O(SHF\Lambda_0) \qquad (4.5)$$

Define $R^X_{FIFO-EA}$ as the running time of the FIFO-EA algorithm for Architecture X (X can be M or T). For Architecture M,

$$R^M_{FIFO-EA} = O(S\bar{H}F\Lambda_{0_M})$$

For Architecture T, because WAN node-pairs are tunneled, there are no gaps between consecutive time segments, and the FIFO-EA Algorithm only need to find the earliest time among the end time $t_s$'s of the last assignment on each wavelength channels. Therefore,

$$R^T_{FIFO-EA} = O(F\Lambda_{0_T})$$

Table 4.1: Running time results of the Shortest-Parth-Routing Algorithm, the Mathematically-Optimal Algorithm, and the FIFO-EA Algorithm. $N_V$ is the number of nodes in the WAN. $N_E$ is the number of edges in the WAN. $S$ is the number of time segments for transactions scheduled over one wavelength on a link ($S \leq N_s$ for Architecture M). $\bar{H}$ is the average number of hops for all paths. $F$ is the number of fibers per link. $\Lambda_{0_M}$ is the number of wavelengths per fiber for Architecture M, and $\Lambda_{0_T} = \Lambda_{0_M}/D$ is the number of wavelengths per tunnel per fiber for Architecture T.

| Running Time | Architecture M | Architecture T |
|:---:|:---:|:---:|
| SPR | $O(N_V^3)$ | $O(1)$ |
| MO | $O(2^{N_E}F\Lambda_{0_M})$ | $O(F\Lambda_{0_T})$ |
| FIFO-EA | $O(S\bar{H}F\Lambda_{0_M})$ | $O(F\Lambda_{0_T})$ |

**Analysis of Coefficients of the Running Time**

As discussed in [16, Chap.25.2], the constant hidden in the $\Theta$-notation for the running time of the Floyd-Warshall Algorithm, $\Theta(N_V^3)$, is small. If only considering comparison, summation, and assignment operations, the coefficient of $\Theta(N_V^3)$ for the Floyd-Warshall Algorithm, including both computing the shortest-path weights and constructing the shortest path, is six.

In the derivation of the running time of the FIFO-EA Algorithm, the omitted constant in $O(SHF\Lambda)$ is three, and, $2SHF\Lambda < T_1(S) < 3SHF\Lambda$. During the execution of the worst case of algorithm FIFO-EA, for each time segment, there are four instructions involved, storing, loading, comparison, and discarding; as a result, $R^X_{FIFO-EA} \le 4T_1(S)$.

## 4.6.2 Processing Power Analysis of the FIFO-EA Algorithm

The SPR algorithm is run only when there is a change in the network topology, much less frequently than the FIFO-EA algorithm. Hence, we focus on the processing power requirements of the FIFO-EA algorithm.

As defined in Chap.3, $N_V$ is the number of nodes in the WAN, $\rho$ is the average loading factor of each link, and $\Delta$ is the average node degree in the WAN. Let $N_p$ be the average population size connected to one WAN node, $\lambda$ be the request arrival rate at end user, and $\bar{T}_d$ be the average transmission time. Then the required processing power for Architecture X is $\lambda N_p N_V R^X_{FIFO-EA}$ for centralized scheduling, and $\lambda N_p R^X_{FIFO-EA}$ for distributed scheduling. For $\rho$ less than one, from Little's Theorem,

$$\rho = \frac{\lambda N_p \bar{T}_d}{\Delta F \Lambda_0} \tag{4.6}$$

Therefore,

$$\lambda N_p = \frac{\rho \Delta F \Lambda_0}{\bar{T}_d} \tag{4.7}$$

The total processing power for Architecture X can then be re-expressed as, with centralized scheduling,

$$P^X_{cen} = \frac{\rho \Delta F \Lambda_0 N_V R^X_{FIFO-EA}}{\bar{T}_d} \tag{4.8}$$

and, with distributed scheduling,

$$P^X_{dis} = \frac{\rho \Delta F \Lambda_0 R^X_{FIFO-EA}}{\bar{T}_d} \tag{4.9}$$

**Traffic Model**

To obtain numerical values of $P^X_{cen}$ and $P^X_{dis}$, we refer to the traffic model in [26], where the authors estimated the near-future traffic pattern based on the projected US population, and the assumptions that a user is active on the network with probably 0.1, and 10% of the time when a user is active, he/she is sending or receiving data with

81

rate 100 Mbps. From their results, the estimated traffic from one WAN node (Chicago for example) takes up 2526 full wavelengths of 10 Gbps, based on the projected US population and network usage. Therefore, $\rho \Delta F \Lambda_0 \approx 2600$. Furthermore, we choose $\rho = 0.7$, $N_s/F\Lambda_0 = 2$, $\Delta = 2.6$, and $\bar{T}_d = 1s$. From Fig. (3-10), the blocking probabilities of both Architecture M and T with the above parameters are less than $10^{-3}$. From Eq. (3.6) and (3.15), we obtain the average integer values of $S$ for Architecture M and T are both one.

## Average Processing Power

The above traffic model values refer to the network steady state values. Substituting them into Eq. (4.8) and (4.9) we can obtain the average processing power required for the FIFO-EA Algorithm. Substituting $R_{FIFO-EA}^X$ with the running time of the Mathematically-Optimal Algorithm, we can also obtain the average processing power for the Mathematically-Optimal Algorithm. The numerical values are summarized in Table 4.2.

## Peak Processing Power

It is a common engineering practice to design a system to be able to sustain not only the average demand, but the peak demand, which corresponds to the peak traffic arrivals in our problem. Define peak traffic arrivals as the rare case situation where more arrivals than average arrive during 1 ms[5] with the probability of occurrence $< 10^{-3}$.[6] Define $N$ to be a Poisson random variable with mean as the average number of traffic arrivals within one 1 ms from all users in a MAN. With the average traffic arrival rate, within 1 ms, the number of average traffic arrivals in 1 ms is $\bar{N} = \lambda N_p \cdot 1ms = 2.6$. Define $N_{\text{peak}}$ to be the minimum number of peak arrivals within 1 ms, such that,

$$P\left(N > N_{\text{peak}}\right) < 10^{-3}$$

and $N_{\text{peak}} = 9$. Letting $\rho \Delta F \Lambda_0 = 9$ and $\bar{T}_d = 1$ ms in Eq. (4.8) and (4.9), we can obtain the peak processing power for both the FIFO-EA Algorithm and the Mathematically-Optimal Algorithm, which are summarized in Table 4.3.

---

[5]1 ms is chosen because processing of a request should be finished in $\sim$1 ms.

[6]Traffic arrivals more than peak traffic arrivals are discarded, and the blocking probability due to more traffic arrivals than peak is less than $10^{-3}$, which is a reasonable target blocking probability.

Table 4.2: Average processing power for the FIFO-EA Algorithm and the Mathematically-Optimal (MO) Algorithm per Scheduler[*].

| Average Processing Power | FIFO-EA | | MO | |
|---|---|---|---|---|
| | Meshed | Tunneled | Meshed | Tunneled |
| **Centralized Scheduling** | 10.7 GIPS[†] | 29.1 MIPS[††] | $3.5 \times 10^{22}$ GIPS | 29.1 MIPS |
| **Distributed Scheduling** | 178 MIPS | 0.48 MIPS | $5.5 \times 10^{20}$ GIPS | 0.48 MIPS |

[*] The values in this table is the processing power per Scheduler. With centralized scheduling, the values are also the total processing power for scheduling. However, with distributed scheduling, the table values need to be multiplied by $N_V$ to obtain the total processing power of all Schedulers for scheduling.

[†] GIPS is short for Giga Instruction per Second.

[††] MIPS is short for Mega Instructions per Second.

Table 4.3: Peak processing power for the FIFO-EA Algorithm and the Mathematically-Optimal (MO) Algorithm per Scheduler[*].

| Peak Processing Power | FIFO-EA | | MO | |
|---|---|---|---|---|
| | Meshed | Tunneled | Meshed | Tunneled |
| **Centralized Scheduling** | 37.2 GIPS[†] | 0.1 GIPS | $1.2 \times 10^{23}$ GIPS | 0.1 GIPS |
| **Distributed Scheduling** | 0.6 GIPS | 1.7 MIPS[††] | $1.9 \times 10^{21}$ GIPS | 1.7 MIPS |

[*] The values in this table is the processing power per Scheduler. With centralized scheduling, the values are also the total processing power for scheduling. However, with distributed scheduling, the table values need to be multiplied by $N_V$ to obtain the total processing power of all Schedulers for scheduling.

[†] GIPS is short for Giga Instruction per Second.

[††] MIPS is short for Mega Instructions per Second.

From Table 4.3, we observe that by tunneling the WAN node pairs, the peak processing power per Scheduler by the FIFO-EA Algorithm can be reduced by three orders of magnitude for both centralized and distributed scheduling. Taking the processing power of the widely available Intel core processor for PC, i7 Extreme Edition 3960X (Hex core) as a reference, it can perform 177.73 GIPS at 3.33 GHz. The FIFO-EA Algorithm with either centralized or distributed scheduling can be handled by one such CPU per Scheduler for both Architecture M and T. However, with Architecture M, the Mathematically-Optimal Algorithm for distributed scheduling takes a super computer 12 mins to compute the schedule for one request, while with Architecture T, the Mathematically-Optimal Algorithm can be handled by one i7 CPU per Scheduler for both centralized and distributed scheduling.

# 4.7 Control Traffic Analysis of the Scheduling of Optical Flow Switching

The control traffic analysis for the Scheduling of Optical Flow Switching concerns the control traffic involved for users to request for a connection, for command exchange between a Scheduler and switches to set up or tear down a connection. The control traffic analysis applies to both optimal scheduling algorithms and heuristics algorithms.

## 4.7.1 Centralized Scheduling

In **centralized scheduling**, all requests are sent to one Scheduler and the Scheduler decides when and along which lightpath a transmission starts. Therefore, the Scheduler knows every transactions in the network (assuming there is no failure in the network). Reconfiguration commands are sent from the Scheduler to switches and the source and destination nodes that need to be reconfigured to set up an end-to-end lightpath. For a lightpath with $\bar{H}_{dp}$ hops[7], there are $\bar{H}_{dp} - 1$ switches that need to be reconfigured. The request for connection is in the form of $\langle v_s, v_d, t_d \rangle$, where $v_s$ and $v_d$ are the 128-bit IPv6 addresses of the source and destination nodes. The reconfiguration command is in the form of $\langle fiber_{in}, fiber_{out}, \omega, T_s, T_e \rangle$, where $\omega$ specifies the wavelength of the lightpath,

---

[7]Subscripts $dp$ and $cp$ are used to differentiate $\bar{H}$ of data plane and $\bar{H}$ of control plane. Note that for Architecture M, $\bar{H}_{dp} = \bar{H}_{cp}$ and both equal to the average hop number of all paths over the network logical topology, while for Architecture T, $\bar{H}_{dp} = 1$ and $\bar{H}_{cp}$ equals to the average hop number of all paths over the network logical topology.

$T_s$ and $T_e$ specify the start and end time of the transmission.

Assume control traffic is carried over the Internet using TCP/IPv6[8]. The switch reconfiguration command is sent from the Scheduler to the corresponding switch using one TCP session. In one TCP session, at least eight IP packets are sent across the network: three for TCP initialization, three for TCP session close-down, one for the command to the switch, and one for the ACK from the switch to the Scheduler. For TCP session initialization and session close-down, no payload is carried in the IP packets, therefore, the minimum IP packet size is 40 bytes (the header size). To carry the data $\langle fiber_{in}, fiber_{out}, \omega, T_s, T_e \rangle$, $(2L_F + L_{\Lambda_0} + |T_s| + |T_e|)(1 + \alpha_{\text{FEC}})$ bits are needed, where $L_F$ and $L_{\Lambda_0}$ are the number of bits taken to specify a fiber and a wavelength, respectively, and $\alpha_{\text{FEC}}$ denotes the percentage of overhead from the Forward Error Correction relative to the actual length of payload. Therefore, the minimum size for an IP packet carrying schedule data is $320 + (2L_F + L_{\Lambda_0} + |T_s| + |T_e|)(1 + \alpha_{\text{FEC}})$ bits. Define $L_{sch}^c$ to be the total control traffic for one TCP session with reconfiguration or schedule data, then

$$L_{sch}^c = (2L_F + L_{\Lambda_0} + |T_s| + |T_e|)(1 + \alpha_{\text{FEC}}) + 320 \cdot 8 \text{ bits} \qquad (4.10)$$

where $320 \cdot 8$ bits accounts for the bits of IPv6 headers of eight IP packets in one TCP session. Similarly, the total control traffic for one TCP session for a request from the source node to the Scheduler is

$$L_{req}^c = (|t_d| + 128 \cdot 2)(1 + \alpha_{\text{FEC}}) + 320 \cdot 8 \text{ bits} \qquad (4.11)$$

$L_{sch}^c$ and $L_{req}^c$ are the same for both Architecture M and T. Note that after the schedule command is sent to a switch, when the corresponding configuration is set up or torn down, the switch informs the Scheduler that the required configurations have been carried out successfully. Therefore, the total amount of control traffic involved in centralized scheduling of one request is

$$L_{req-sch}^{cX} = L_{req}^c + 3 \cdot (\bar{H}_{dp}^X + 1) \cdot L_{sch}^c \qquad (4.12)$$

where the multiplicative factor 3 accounts the traffic for sending the configuration command from the Scheduler to the switch after the scheduling computation, and the

---

[8]IPv6 is used for future-compatibility.

feedback traffic from a switch to the Scheduler after a connection is set up and torn down. The superscript $X$ can be either $M$ or $T$ to indicate the specific Architecture M or T under study.

Therefore, the total amount of control traffic in centralized scheduling, $L_{cen}^X$, is

$\lambda N_p N_V \bar{H}_{cp} L_{req-sch}^c$, and with Eq. (4.6),

$$L_{cen}^X = \frac{\rho \Delta F \Lambda_0 N_V \bar{H}_{cp}^X L_{req-sch}^{cX}}{\bar{T}_d} \qquad (4.13)$$

To obtain numerical values of $L_{cen}^X$, we refer to the same traffic model in Section 4.6.2 and network topology parameters in Section 3.4 in Chap.3. Table 4.4 lists the parameters we will use to calculate control traffic for both Architecture M and Architecture T. For Architecture T, because the logical data network topology is tunneled, $\bar{H}_{dp}^T = 1$ and the other parameters are the same as in Architecture M[9]. Substituting the values in Table 4.4 into Eq. (4.10), (4.11), and (4.12), we obtain the per-session control traffic values for Architecture M and Architecture T in Table 4.5.

---

[9]Note that $\bar{D}$, $F$ and $\Lambda$ for Architecture M and T are different. However, since they use the same physical topology, $\bar{D}F\Lambda$ is the same for these two different architectures, and the difference in $F$ and $\Lambda$ between these two architectures does not affect the value of $L_{sch}^c$ since a one-byte bit string is used to label different fibers and wavelengths.

Table 4.4: Numerical values of Architecture M and Architecture T for control traffic volume calculation.

| Parameter | Value |
|---|---|
| $\rho\bar{D}F\Lambda_0/\bar{T}_d$ | 2600 for the case of average traffic arrivals. |
| $\rho\bar{D}F\Lambda_0/\bar{T}_d$ | 9000 for the case of peak traffic arrivals[†] |
| $L_F$ | 8 bits |
| $L_{\Lambda_0}$ | 8 bits |
| $N_V$ | 60 |
| $\bar{H}_{dp}^M$ | 4 |
| $\bar{H}_{cp}^M$ | 4 |
| $\bar{H}_{dp}^T$ | 1 |
| $\bar{H}_{cp}^T$ | 4 |
| $|T_s|$ and $|T_e|$ | 80 bits. The Oracle timestamp format with fractional seconds of 6 digit precision (mili-second) is 10 bytes |
| $|t_d|$ | 64 bits. $t_d$ can be expressed as an integer value of unit mili-second using Int64 format which is 8 bytes. 24 hours can be stored using a 37-bit integer. Therefore, using a 64-bit integer, a time duration of 910117 years can be stored. |
| $\alpha_{\text{FEC}}$ | 10% |

[†] Peak traffic arrivals is defined in Section 4.6.2 as the rare case situation where more arrivals than average arrive during 1 ms with the probability of occurrence $< 10^{-3}$.

Since under the centralized scheduling control, the scheduler knows the configuration of the whole network, and no link state update on connection setups and teardowns is necessary. Therefore, the control traffic corresponding to link state update is the same as the scheduling control traffic for both Architecture M and T, and hence, we do not need to calculate the link state traffic separately under centralized scheduling

Table 4.5: Numerical values for per-session control traffic under centralized scheduling for both Architecture M and Architecture T.

| Parameter | Value |
|---|---|
| $L_{sch}^c$ | 2763 bits |
| $L_{req}^c$ | 2912 bits |
| $L_{req-sch}^{cM}$ | 44357 bits |
| $L_{req-sch}^{cT}$ | 19490 bits |

control.

## 4.7.2  Distributed Scheduling

In the **distributed scheduling** with Architecture M, when a connection request arrives at a node, the node runs the scheduling algorithm locally, and broadcasts the scheduling result to the whole network immediately, so that every other node gets an updated view of the whole network. Therefore, switch reconfiguration control traffic can be replaced by link state update control traffic. In this case, $L_{req}^d$, the control traffic for requests, is the same as $L_{req}^c$; and, $L_{sch}^d$, the control traffic for scheduling to be sent to one other node is the same as $L_{sch}^c$. The amount of control traffic involved in the distributed scheduling for one request, $L_{req-sch}^{dX}$, is

$$L_{req-sch}^{dM} = L_{req}^d + N_V \cdot 3 \cdot L_{sch}^d \text{ for Architecture M}$$

The total amount of control traffic in Architecture M is

$$L_{dis}^M = \frac{\rho \Delta F \Lambda_0 N_V \bar{H}_{cp}^M \cdot L_{req-sch}^{dM}}{\bar{T}_d}$$

In Architecture T since WAN node pairs are tunneled, the scheduling information of one request only needs to be broadcast to the source and destination nodes, and,

$$L_{req-sch}^{dT} = L_{req}^d + 3 \cdot 2L_{sch}^d \text{ for Architecture T}$$

88

The total amount of control traffic in Architecture T is

$$L_{dis}^T = \frac{\rho \Delta F \Lambda_0 N_V \bar{H}_{cp}^T \cdot L_{req-sch}^{dT}}{\bar{T}_d}$$

Similar as the analysis in Section 4.6.2, we use numerical values in Table 4.4 to calculate control traffic for both the average case ($\rho \bar{D} F \Lambda_0 / \bar{T}_d = 2600$) and the case of peak traffic arrivals ($\rho \bar{D} F \Lambda_0 / \bar{T}_d = 9000$). Table 4.6 summarizes the average control traffic for Architecture M and T under both centralized scheduling and distributed scheduling, while Table 4.7 summarizes the peak control traffic. To better understand the control traffic burden onto the control plane, we break down the control traffic in Table 4.6 and 4.7 into average edge load and heaviest edge load. The average bidirectional load per edge is obtained by dividing the total traffic by the total number of bidirectional edges in the WAN, which is 77. For the centralized scheduling, the control traffic happens between the central Scheduler with users or switches. Consider the central Scheduler has an average node degree 2.6, then the load on the edges connecting the central Scheduler is the heaviest, and can be obtained from $L_{cen}^X / (\bar{H}_{cp}^X \bar{\Delta})$. For distributed scheduling, the traffic is between all schedulers with users and switches, with the assumption of uniform all-to-all traffic and a symmetric network topology, the average bidirectional load per edge is also the heaviest edge load.

Considering the peak control traffic on the heaviest-edge in Table 4.7 with distributed scheduling, tunneling leads to a two orders of magnitude reduction of the heaviest-edge control traffic.

Table 4.6: Average total control traffic of both centralized scheduling and distributed scheduling for both Architecture M and Architecture T.

| Average Control Traffic | | Architecture M | Architecture T |
|---|---|---|---|
| **Centralized Scheduling** | Total traffic | 27.68 Gbps | 12.16 Gbps |
| | Per-edge | 0.36 Gbps | 0.16 Gbps |
| | Heaviest-edge | 2.66 Gbps | 1.17 Gbps |
| **Distributed Scheduling** | Total traffic | 312.16 Gbps | 12.16 Gbps |
| | Per-edge[†] | 4.05 Gbps | 0.16 Gbps |

[†] With distributed scheduling, in a symmetric WAN topology using shortest path routing, all control traffic is distributed evenly to all links. Therefore, the average bidirectional load per edge is also the heaviest edge load with distributed scheduling.

Table 4.7: Peak total control traffic of both centralized scheduling and distributed scheduling for both Architecture M and Architecture T.

| Peak Control Traffic | | Architecture M | Architecture T |
|---|---|---|---|
| **Centralized Scheduling** | Total traffic | 95.8 Gbps | 42.1 Gbps |
| | Per-edge | 1.24 Gbps | 0.55 Gbps |
| | Heaviest-edge | 9.2 Gbps | 4.1 Gbps |
| **Distributed Scheduling** | Total traffic | 1.08 Tbps | 42.1 Gbps |
| | Per-edge[†] | 14 Gbps | 0.54 Gbps |

[†] With distributed scheduling, in a symmetric WAN topology using shortest path routing, all control traffic is distributed evenly to all links. Therefore, the average bidirectional load per edge is also the heaviest edge load with distributed scheduling.

# Chapter 5

# Physical Layer Impairments in EDFA-Amplified OFS Mesh Networks

Current optical networks are operated with predominantly static connections, and lightpaths are changed quasi-statically and may remain unchanged for months. Present methods of setting up a wavelength path result in slow changes to the network (~20 min setups), as each of the network element along the path is gradually tuned to the final settings to avoid instabilities arising from rapidly introducing another optical channel into the network. The link quality of all adjacent links are monitored as the lightpath is turned on in several steps. Optical Flow Switching enables and coordinates per-session end-to-end all optical lightpath switching over an optical network of meshed physical topology. Irrespective to whether the logical topology is Meshed or Tunneled, over the meshed physical topology, merging and diverging traffic in any link, unless the MAN owns an entire fiber tunnel, is unavoidable. Therefore, lightpaths for different source-destination pairs may share common mid-span fiber links, and there are dynamic lightpath-switching over the network, both in the aggregation sites in the MAN and in the long-haul backbone network.

In a long-haul fiber connection in today's optical networks, one common way of amplifying the optical signal before it degrades is to use EDFAs. Pumped by an out-band laser, an EDFA provides amplification to all wavelength channels in-band. To mitigate the long-distance fiber loss in the long-haul network and to avoid frequent regeneration of optical signals, a tandem of 10's of EDFAs can be used for amplification

of the optical signals. However, due to both the fundamental limitations and the gain control of EDFAs, when lightpath is switched on/off dynamically in a meshed network, existing channels in the same fiber experience fast transients and cross-channel power coupling, which is then amplified and controlled by downstream EDFAs, further deteriorating the problem. In [35, 45, 36, 29, 30], the authors have investigated cross-channel power transients and excursions under constant gain control of EDFA in optical networks. However, these work has only investigated the effects on average optical power of each channel, which is not enough to analyze the impact on the channel quality. The pertinent metric for channel quality is Bit Error Rate (BER), which is a result of the detection of the corrupted signal in random noise from the EDFAs and the receiver. In the experiment by Junio [23] done at Bell Lab, fast transients and steady-state channel quality variations were examined. The former impairment can be quenched using "adiabatic" switching and the new Transport Layer Protocol in [21], and from the scheduling perspective we assume the fast transient is totally accounted for by the legal modulation and the transport layer protocol. The second impairment directly impacts channel quality, and is a result of several factors: the randomness in the EDFA gain, accumulation of Amplified Spontaneous Noise (ASE), constant gain control of EDFAs, and the receiver noise. Switching-induced channel-quality variations affect scheduling, i.e., worst case channel quality during the scheduled transmission time needs to be estimated to guarantee performance and reach (for a selected transmission rate if flexible transmission rate is used).

## 5.1   Detection Model

We assume the signal is On-Off modulated (at 10 Gbits/s in the experiments), and use a matched filter at the receiver to obtain the "sufficient" statistics of the symbol. Let $X$ represent the random variable at the matched filter output. Let Hypothesis "1" ($H_1$) be the case where the received bit is "1", and Hypothesis "0" ($H_0$) be the case where the received bit is "0". The signal fluctuations in the "1" and "0" levels, are randomized by accumulated noise along the link. The accumulated noise and fluctuations of the amplitudes of the signal will be compounded over the cascaded amplifiers. The received output from the matched filter, $X$, can be modeled as a

Gaussian random variable under each hypothesis, and,

$$H_1 : \text{Gaussian}(\mu_s, \sigma_s^2) \to f_{x|H_1}(x) = \frac{1}{\sigma_s\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_s)^2}{2\sigma_s^2}\right]$$
$$H_0 : \text{Gaussian}(\mu_n, \sigma_n^2) \to f_{x|H_0}(x) = \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_n)^2}{2\sigma_n^2}\right] \tag{5.1}$$

Since "1"s and "0"s are assumed to be equally likely, $P(H_1) = P(H_0) = 1/2$. For the detected signal, the variance of bit "1" $\sigma_s^2$ is larger than the variance of bit "0" $\sigma_n^2$, and $\mu_n < \mu_s$. Detection of Gaussian Signals in Gaussian Noise has been extensively studied, e.g. in [37]. Using the same techniques in [37], we re-derived the optimal detector for our problem in Appendix B. Appendix B shows the optimal threshold detector is a two-threshold detector, and under most realistic conditions, can be approximated by a single-threshold detector. We summarize the results in Appx. B in the following two propositions.

**Proposition 1.** *The optimal two-threshold detector to detect the hypothesis in Eq. (5.1) is:*

$$H_1 \ is \ true, \ if \quad X < x_2^*, \ or \ X > x_1^*$$
$$H_0 \ is \ true, \ if \quad x_2^* < X < x_1^* \tag{5.2}$$

*where,*

$$x_1^* = a + \sqrt{a^2 - b}$$
$$x_2^* = a - \sqrt{a^2 - b} \quad, \ where$$
$$a = \frac{\mu_n \sigma_s^2 - \mu_s \sigma_n^2}{\sigma_s^2 - \sigma_n^2} \tag{5.3}$$
$$b = \frac{\sigma_s^2 \mu_n^2 - \sigma_n^2 \mu_s^2 + 2\sigma_s^2 \sigma_n^2 \log\left(\frac{\sigma_n}{\sigma_s}\right)}{\sigma_s^2 - \sigma_n^2}$$

*The minimum error probability with this two-threshold detector is,*

$$(P_e)_{\min} = \frac{1}{2}\left[Q\left(\frac{\mu_s - x_1^*}{\sigma_s}\right) + Q\left(\frac{x_1^* - \mu_n}{\sigma_n}\right)\right] - \frac{1}{2}\left[Q\left(\frac{\mu_s - x_2^*}{\sigma_s}\right) - Q\left(\frac{\mu_n - x_2^*}{\sigma_n}\right)\right] \tag{5.4}$$

*where $Q(x)$ is the Q-function, and,*

$$Q(x) = \frac{1}{2} \, erfc\left(\frac{x}{\sqrt{2}}\right)$$

**Proposition 2.** *If,*

$$(\mu_n \sigma_s^2 - \mu_s \sigma_n^2)^2 \gg \sigma_s^2 \mu_n^2 - \sigma_n^2 \mu_s^2 + 2\sigma_s^2 \sigma_n^2 \log\left(\frac{\sigma_n}{\sigma_s}\right)$$

*then the two-threshold detector in Eq. (5.2) can be approximated by the single-threshold detector below,*

$$H_1 \ \text{is True, if } X > x_1^*$$
$$H_0 \ \text{is True, if } X < x_1^*$$

*where $x_1^*$ is defined in Eq. (5.3). The error probability of this single-threshold detector is,*

$$(P_e)_{single\text{-}threshold} = \frac{1}{2}\left[Q\left(\frac{\mu_s - x_1^*}{\sigma_s}\right) + Q\left(\frac{x_1^* - \mu_n}{\sigma_n}\right)\right] \tag{5.5}$$

In the literature [3, Chap. 4] of studying BER of optical receivers, single-threshold model was often used. The single-threshold model was further simplified by ignoring the term $\log\left(\frac{\sigma_n}{\sigma_s}\right)$ in Eq. (5.3), which made sense if $\sigma_n^2$ and $\sigma_s^2$ were comparable. Note that if $\log\left(\frac{\sigma_n}{\sigma_s}\right) \approx 0$, then,

$$\frac{(x_1^* - \mu_n)^2}{2\sigma_n^2} = \frac{(x_1^* - \mu_s)^2}{2\sigma_s^2}$$

The threshold can be solved explicitly, and is,

$$x_1^* = \frac{\sigma_n \mu_s + \sigma_s \mu_n}{\sigma_s + \sigma_n}$$

Define,

$$Q \triangleq \frac{x_1^* - \mu_n}{\sigma_n} = \frac{\mu_s - x_1^*}{\sigma_s} \tag{5.6}$$

This parameter Q (different from the Q-function in Proposition 1 and Proposition 2) is often used to indicate BER [3, Chap. 4] as

$$BER = \frac{1}{2}\text{erfc}(\frac{Q}{\sqrt{2}}) \tag{5.7}$$

94

However, in our analysis, we consider the situation where, in the detector, the variance of "1" bit $\sigma_s^2$ is much larger than the variance of "0" bit $\sigma_n^2$ due to noise accumulation through multiple EDFAs[1]. Therefore, we can not ignore the term $\log\left(\frac{\sigma_n}{\sigma_s}\right)$, and we can not use Eq. (5.6) and (5.7) for calculating BER.

## 5.2  Noise Sources in EDFA-Amplified Optical Links

### 5.2.1  Amplified Spontaneous Emission

Spontaneous emission is a major noise source in both lasers and amplifiers. It is mostly a result of the random de-excitation of the laser ions, which has been understood through quantum mechanics [20, 43]. In an amplifier, the spontaneous emission noise is further amplified when it travels through the rest of the amplification medium, therefore, it is named Amplified Spontaneous Emission (ASE).

For an optical amplifier with optical bandwidth $B_o$, the ASE noise can be written as,

$$P_{sp} = 2S_{sp}B_o \tag{5.8}$$

where $B_o$ is bandwidth of the optical filter. The parameter $S_{sp}$ is the spectral density of the ASE noise, which can be considered as nearly constant, and can be written as [3],

$$S_{sp} = (G - 1)n_{sp}h\nu \tag{5.9}$$

where $\nu$ is the optical frequency of the incident photons, and $h$ is the Planck constant. The parameter $n_{sp}$ is the spontaneous-emission factor, and

$$n_{sp} = \frac{N_2}{N_2 - N_1} \tag{5.10}$$

where $N_1$ and $N_2$ are the atomic populations for the ground and excited states, respectively. Therefore, the ASE noise in Eq. (5.8) can be re-written as,

$$P_{sp} = 2(G - 1)n_{sp}h\nu B_o = 2(G - 1)\frac{N_2}{N_2 - N_1}h\nu B_o \tag{5.11}$$

---

[1]In the experiment the extinction ratio is not zero (16:1 in dB), but after many amplifiers $\sigma_n^2$ is dominated by accumulated amplifier noise.

## 5.2.2 Detection Noise

The shot noise and thermal noise are two fundamental noise sources that limit the BER of an optical receiver. For incident optical signal that itself fluctuates because of noise accumulation along the transmission, additional noise including beat noise and intensity noise is generated at the receiver [3]. Expressing them in the form of receiver current, the current of a photo receiver can be written as

$$I(t) = I_p + i_s(t) + i_{\text{sig-sp}}(t) + i_{\text{sp-sp}}(t) + i_I(t) + i_T(t) \qquad (5.12)$$

where $i_s(t)$ is the zero-mean current fluctuation from shot noise, $i_{\text{sig-sp}}(t) + i_{\text{sp-sp}}(t)$ is from beat noise, $i_I(t)$ is from incident optical signal intensity noise, and $i_T(t)$ is from Thermal noise. $I_P$ is the average photodiode current, and,

$$I_p = R\bar{P}_{in}$$

where $P_{in}$ is the average incident optical power. The parameter $R$ is the photodetector responsivity, and,

$$R = \frac{\eta q}{h\nu}$$

where $\eta$ is the quantum efficiency of the photo detector.

**Shot Noise**

Shot noise is a manifestation of the fact that the electron generation at the receiver constitutes a random process, leading to current fluctuation in the photodiode. The photodiode current generated in response to a constant optical signal $P_{in}$ can be written as,

$$I(t) = I_p + i_s(t) \qquad (5.13)$$

where $I_p = RP_{in}$ is the average photodiode current and $i_s(t)$ is the current fluctuation due to shot noise. The noise variance of the shot noise is,

$$\sigma_s^2 \triangleq \langle i_s^2(t) \rangle = 2qI_pB_e \qquad (5.14)$$

where $B_e$ is the effective noise bandwidth of the receiver.

## Intensity Noise from EDFA Gain Randomness

Intensity fluctuation of incident optical signal leads to intensity noise in the photodetector. In this analysis, we focus on the intensity fluctuation caused by randomness in the gain of the EDFA because a chain of $k$ amplifiers amplifies the intensity noise exponentially (which will be discussed in later sections). Different from the spontaneous emission which is unpolarized, the randomness in the gain of an EDFA refers to the gain fluctuation which is in-phase of the input signal light, and can be best understood from quantum mechanics [25]. We use $g$ to represent the EDFA gain if the gain is approximated as a constant, and use $G$ to represent the EDFA gain if the gain randomness is considered. We quantify the randomness in the gain using $\sigma_G^2/g^2$, where $\sigma_G^2$ is the variance of $G$, and $g$ is the mean of $G$. Randomness in the EDFA gain leads to intensity fluctuation of the amplified optical signal, which is further amplified by downstream EDFAs. Define $\sigma_{IG}^2$ to be the intensity noise variance of the receiver current resulted from the gain randomness, we have,

$$\sigma_{IG}^2 = R^2 \langle (\Delta P_{in})^2 \rangle = R^2 \bar{P}_{in}^2 \frac{\langle (\Delta P_{in})^2 \rangle}{\bar{P}_{in}^2}$$

We will derive $\sigma_{IG}^2$ as a function of $\sigma_G^2/\mu_G^2$ in the later analysis.

## Beat Noise

At the photo receiver, the generated photocurrent can be written as [3, Chap.6],

$$I = R|\sqrt{G}E_s + E_{sp}|^2 + i_s + i_T \tag{5.15}$$

where $E_s$ is the signal field and $E_{sp}$ is the optical field associated with the ASE. ASE-induced current noise manifests into beat noise, which is a result of beating of $E_s$ with $E_{sp}$ and $E_{sp}$ with itself [3, Chap.6]. The ASE field $E_{sp}$ is broadband and can be written as

$$E_{sp} = \int \sqrt{S_{sp}} \exp(\phi_n - i\omega_n t) d\omega_n$$

where $\phi_n$ is the phase of the noise-spectral component at the frequency $\omega_n$. Using $E_s = \sqrt{P_s} \exp(\phi_s - i\omega_s t)$, the current of the two beating noise terms can be obtained

from Eq. (5.15) as

$$i_{\text{sig-sp}} = 2R \int (gP_sS_{sp})^{1/2} \cos\theta_1 d\omega_n \tag{5.16}$$

$$i_{\text{sp-sp}} = 2R \int \int S_{sp} \cos\theta_2 d\omega_n d\omega_n' \tag{5.17}$$

where $\theta_1 = (\omega_s - \omega_n)t + \phi_n - \phi_s$ and $\theta_2 = (\omega_n - \omega_n')t + \phi_n' - \phi_n$ are two random phases that vary rapidly. Averaging over the random phases, we obtain the noise variance of the two beat noise,

$$\sigma_{\text{sig-sp}}^2 = 4R^2 gP_sS_{sp}B_e$$

$$\sigma_{\text{sp-sp}}^2 = 4R^2 S_{sp}^2 B_o B_e$$

**Thermal Noise**

Thermal noise is a manifestation of random thermal motion of electrons in any conductor. The noise variance due to thermal noise can be written as

$$\sigma_T^2 \triangleq \langle i_T^2(t) \rangle = \frac{4k_BT}{R_L}B_e$$

where $k_B$ is the Boltzmann constant, $T$ is the absolute temperature, and $R_L$ is the load resistor. Thermal noise is small relative to the other noise sources and are often ignored in detection models.

## 5.3 EDFA Architecture

Figure 5-1 shows a typical EDFA architecture. This is also the EDFA used in the experiments in [23]. It is a two-stage amplifier with each span of Erbium-Doped Fibers pumped by a 980 nm source. A Gain-Flattening Filter (GFF) is used between them to flatten the gain spectrum. A control system (controlling the mid-stage Variable-Optical Attenuator between the two-stage Erbium-Doped fibers and the pump source) is capable of adjusting the internal gain to within 0.1 dB of the target within milliseconds. Define $P_{\text{in}}$ as the input optical power to an EDFA, and $P_{\text{out}}$ as the output optical power, and $g = P_{\text{out}}/P_{\text{in}}$ as the average gain. In practice, there are two common operating modes for EDFA: constant gain mode, and constant power mode. In the constant power mode, the amplifier is controlled to produce a constant output

power, regardless of input power variations. In the constant gain mode, the amplifier is controlled to maintain the average of $P_{\text{out}}/P_{\text{in}}$ to be constant. Denote $P_{sp}$ as the noise of the amplified spontaneous emission (ASE) that is added to the input signal by the EDFA. Under the constant gain control, the ratio of the total output power over the total input power (i.e., $(P_{\text{in}} + P_{\text{sp}})/P_{\text{in}}$, instead of the amplification gain seen by input signal) is controlled to be constant. To avoid misunderstanding, we use quotation signs for the constant gain control of the ratio of the total output power to the input power, i.e., "constant gain mode" or "constant gain control". Figure 5-2 shows a simplified schematic representation of and EDFA with random gain $G$ and ASE noise $P_{sp}$.



EDF: Erbium-Doped Fiber                VOA: Variable Optical Attenuator
WDM: Wavelength Division Multiplexer    GFF: Gain Flatten Filter
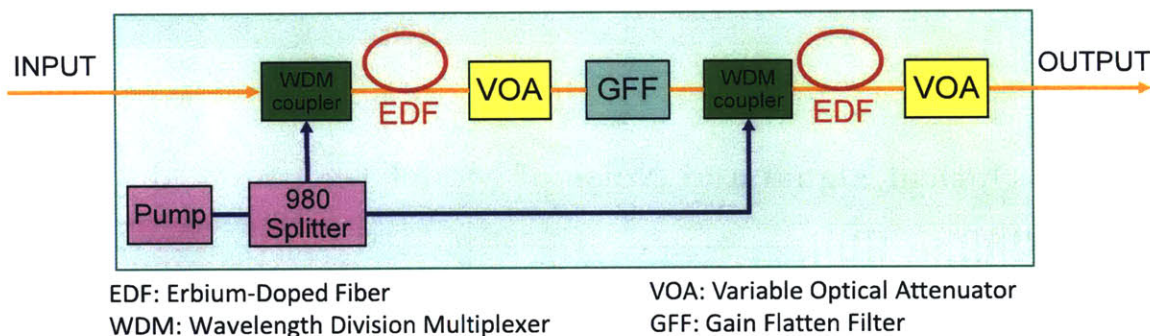
Figure 5-1: The architecture of the EDFA used for the experiments. It is a two-stage amplifier pumped by a 980 nm pump source. The variable optical attenuator and the pump can be adjusted by the EDFA control system.
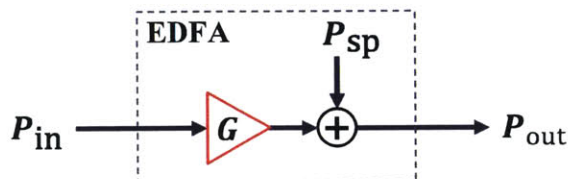


Figure 5-2: A schema of an EDFA with random gain $G$ and ASE noise $P_{sp}$.

Define a naming convention, Model A-B, for EDFA-amplified optical links, where,

- $A$ refers to the number of cascaded EDFA(s) in the link,

- $B = g$ if the amplifier gain intensity noise is ignored, and $B = G$ if otherwise.

## 5.4   EDFA-Amplified Optical Link Model 1-g

We start with analyzing the simplest EDFA model, Model 1-g as shown in Fig. 5-3. In the model, an optical attenuation $l$ is applied before the EDFA to simulate the loss

the optical signal experiences when propagating in the fiber before being amplified by an amplifier.
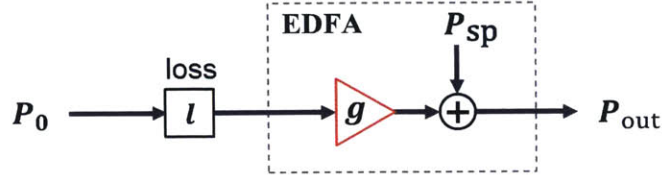


Figure 5-3: Model 1-g with fiber loss $l$.

## 5.4.1 Optical Signal and Noise of Model 1-g

The output $P_{\text{out}}$ includes both signal and noise, and

$$P_{\text{out}} = P_0 lg + P_{sp}$$

where $P_0$ is the initial optical signal power out of a transmitter. $P_{sp}$ is the optical ASE noise from Eq. (5.11), and $P_{sp} = 2(g-1)n_{sp}h\nu B_o$. The optical signal to noise ratio (OSNR) at the output, therefore, is

$$\text{OSNR}_{1-g} = \frac{P_0 lg}{2(g-1)n_{sp}h\nu B_o}$$

## 5.4.2 Electrical Signal and Noise of Model 1-g

More noise (shot noise and thermal noise) is added to the signal at the optical receiver when being detected, and ASE noise manifests into beat noise both with the signal field and with itself. Assuming infinite extinction ratio for the On-Off Keying modulation and ignoring thermal noise, from Eq. (5.15), the photocurrent is,

$$I_1 = R|E_0\sqrt{lg} + E_{sp}|^2 + i_{s_1} + i_{\text{sig-sp}_1} + i_{\text{sp-sp}_1}, \quad \text{for bit ``1''}$$
$$I_0 = R|E_{sp}|^2 + i_{s_0} + i_{\text{sp-sp}_0}, \quad \text{for bit ``0''}$$

100

Therefore, for bit "1", we have,

$$\bar{I}_1 = R(P_0 lg + P_{sp})$$

$$\sigma_{s_1}^2 = 2qR(P_0 lg + P_{sp})B_e$$

$$\sigma_{sig-sp_1}^2 = 4R^2 P_0 lg(g-1)n_{sp}h\nu B_e$$

$$\sigma_{sp-sp_1}^2 = 4R^2\left[(g-1)n_{sp}h\nu\right]^2 B_o B_e$$

For bit "0", we have

$$\bar{I}_0 = RP_{sp}$$

$$\sigma_{s_0}^2 = 2qRP_{sp}B_e$$

$$\sigma_{sp-sp_0}^2 = 4R^2\left[(g-1)n_{sp}h\nu\right]^2 B_o B_e$$

The variance of bit "1" is larger than the variance of bit "0". Detecting whether the received bit is "1" or "0" is different from classical formulation where variances of bit "1" and "0" are approximately the same. In [3, Chap.6], when detection noise is dominated by the signal-spontaneous beat noise, the electrical SNR is defined as $\mathrm{SNR}_{el} = \frac{(\bar{I}_1 - \bar{I}_0)^2}{\sigma_{sig-sp}^2}$ However, in our formulation, accumulated noise increases quickly after a cascade of multiple amplifiers, and the accumulated noise power can be large enough so that other noise factors besides signal-spontaneous beat noise also become significant and can not be ignored. Therefore, we define the electrical SNR (ESNR) to be,

$$\mathrm{ESNR} \triangleq \frac{\left(\bar{I}_1 - \bar{I}_0\right)^2}{\sigma_{s_1}^2 + \sigma_{sig-sp_1}^2 + \sigma_{sp-sp_1}^2} \tag{5.18}$$

For Model 1-g,

$$\mathrm{ESNR}_{1-g} = \frac{(P_0 lg)^2}{\frac{2q(P_0 lg + P_{sp})B_e}{R} + 4P_0 lg(g-1)n_{sp}h\nu B_e + 4\left[(g-1)n_{sp}h\nu\right]^2 B_o B_e}$$

Ignoring the shot noise term,

$$\mathrm{ESNR}_{1-g} \approx \frac{P_0 lg}{2P_{sp} + \frac{P_{sp}^2}{P_0 lg}} \cdot \frac{B_o}{B_e}$$

If $P_0 lg \gg P_{sp}$,

$$\mathrm{ESNR}_{1-g} \approx \frac{B_o}{2B_e} \cdot \frac{P_0 lg}{P_{sp}} = \frac{B_o}{2B_e}\mathrm{OSNR}_{1-g}$$

101

Note that the BER defined in Eq. (5.4) or (5.5) depends on statistics of both bit "1" and bit "0", while the ESNR defined in Eq. (5.18) is a function of only bit "1" statistics. Therefore, the ESNR itself is not enough to determine the BER. We included the analysis of the afore-defined ESNR in this thesis because this was similar to the $\sigma_s^2/\mu_s^2$ measured in the experiment by Junio [23].

## 5.5 EDFA-Amplified Optical Link Model k-g

Figure 5-4 shows a cascade of $k$ amplifiers, each separated by a fiber span with loss $l$.
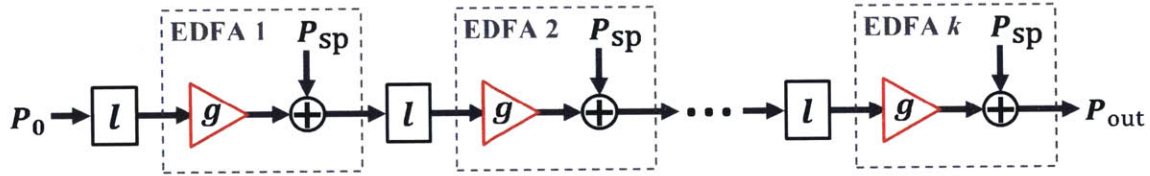


Figure 5-4: Model k-g.

### 5.5.1 Optical Signal and Noise of Model k-g

The optical signal power after $k$ amplifiers $(P_{\text{sig}})_k$ is

$$(P_{\text{sig}})_k = P_0(lg)^k$$

The accumulated ASE noise after $k$ amplifiers $(P_{sp})_k$ is

$$(P_{sp})_k = P_{sp} \sum_{i=0}^{k-1} (lg)^i = \begin{cases} kP_{sp} & \text{if } lg = 1 \\ P_{sp}\frac{1-(lg)^k}{1-lg} & \text{otherwise} \end{cases}$$

The optical signal to noise ratio of Model k-g, therefore, is

$$(\text{OSNR})_{k-g} = \begin{cases} \frac{P_0}{2k(g-1)n_{sp}h\nu B_o} & \text{if } lg = 1 \\ \frac{P_0(lg)^k(1-lg)}{2(g-1)n_{sp}h\nu B_o\left[1-(lg)^k\right]} & \text{otherwise} \end{cases} \tag{5.19}$$

From the above equation, we have the following observations:

1. For $lg < 1$, as $k$ increases, $(\text{OSNR})_{k-g}$ decreases at a rate of $\frac{(lg)^k}{1-(lg)^k}$. And for large $k$, $(\text{OSNR})_{k-g}$ decreases at a rate of $(lg)^k$ asymptotically.

102

- $\lim_{k \to \infty}(\text{OSNR})_{k-g} = 0$

2. For $lg = 1$, as $k$ increases, $(\text{OSNR})_{k-g}$ decreases at a rate of $\frac{1}{k}$.

- $\lim_{k \to \infty}(\text{OSNR})_{k-g} = 0$

3. For $lg > 1$, as $k$ increases, $(\text{OSNR})_{k-g}$ decreases at a rate of $\frac{1}{1 - \frac{1}{(lg)^k}}$.

- $\lim_{k \to \infty}(\text{OSNR})_{k-g} = \frac{P_0(lg-1)}{2(g-1)n_{sp}h\nu B_o}$

## 5.5.2 Electrical Signal and Noise of Model k-g

The photocurrent of bit "1" is,

$$
(I_1)_k = \begin{cases}
R\left|E_0 + E_{sp}\sqrt{k}\right|^2 + i_{s_1} + i_{\text{sig-sp}_1} + i_{\text{sp-sp}_1} & \text{if } lg = 1 \\
R\left|E_0\sqrt{(lg)^k} + E_{sp}\sqrt{\frac{1-(lg)^k}{1-lg}}\right|^2 + i_{s_1} + i_{\text{sig-sp}_1} + i_{\text{sp-sp}_1} & \text{otherwise}
\end{cases}
$$

The photocurrent of bit "0" is,

$$
(I_0)_k = \begin{cases}
Rk|E_{sp}|^2 + i_{s_0} + i_{\text{sp-sp}_0} & \text{if } lg = 1 \\
R\frac{1-(lg)^k}{1-lg}|E_{sp}|^2 + i_{s_0} + i_{\text{sp-sp}_0} & \text{otherwise}
\end{cases}
$$

Therefore, for bit "1", we have,

$$
(\bar{I}_1)_k = \begin{cases}
R(P_0 + kP_{sp}) & \text{if } lg = 1 \\
R\left[P_0(lg)^k + P_{sp}\frac{1-(lg)^k}{1-lg}\right] & \text{otherwise}
\end{cases}
$$

$$
(\sigma_{s_1}^2)_k = \begin{cases}
2qR(P_0 + kP_{sp})B_e & \text{if } lg = 1 \\
2qR\left[P_0(lg)^k + P_{sp}\frac{1-(lg)^k}{1-lg}\right]B_e & \text{otherwise}
\end{cases}
$$

$$
(\sigma_{sig-sp_1}^2)_k = \begin{cases}
4R^2 P_0 k(g-1)n_{sp}h\nu B_e & \text{if } lg = 1 \\
4R^2 P_0(g-1)n_{sp}h\nu B_e\frac{(lg)^k\left[1-(lg)^k\right]}{1-lg} & \text{otherwise}
\end{cases} \tag{5.20}
$$

$$
(\sigma_{sp-sp_1}^2)_k = \begin{cases}
4R^2\left[k(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{if } lg = 1 \\
4R^2\left[\frac{1-(lg)^k}{1-lg}(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{otherwise}
\end{cases} \tag{5.21}
$$

For bit "0", we have

$$(\bar{I}_0)_k = \begin{cases} kRP_{sp} & \text{if } lg = 1 \\ RP_{sp}\frac{1-(lg)^k}{1-lg} & \text{otherwise} \end{cases}$$

$$(\sigma_{s0}^2)_k = \begin{cases} 2kqRP_{sp}B_e & \text{if } lg = 1 \\ 2qRP_{sp}B_e\frac{1-(lg)^k}{1-lg} & \text{otherwise} \end{cases}$$

$$(\sigma_{sp-sp0}^2)_k = \begin{cases} 4R^2\left[k(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{if } lg = 1 \\ 4R^2\left[\frac{1-(lg)^k}{1-lg}(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{otherwise} \end{cases}$$

Therefore, the ESNR as defined in Eq. (5.18) for Model k-g is

$$\text{ESNR}_{k-g} = \begin{cases} \dfrac{P_0^2}{\frac{2q(P_0+kP_{sp})B_e}{R}+4P_0k(g-1)n_{sp}h\nu B_e+4[k(g-1)n_{sp}h\nu]^2 B_o B_e} & \text{if } lg = 1 \\[4ex] \dfrac{\left[P_0(lg)^k\right]^2}{\frac{2q\left[P_0(lg)^k+P_{sp}\frac{1-(lg)^k}{1-lg}\right]B_e}{R}+4P_0(g-1)n_{sp}h\nu B_e\frac{(lg)^k[1-(lg)^k]}{1-lg}+4\left[\frac{1-(lg)^k}{1-lg}(g-1)n_{sp}h\nu\right]^2 B_o B_e} & \text{otherwise} \end{cases}$$

(5.22)

Ignoring the shot noise,

$$\text{ESNR}_{k-g} \approx \begin{cases} \dfrac{P_0}{2kP_{sp}+k^2\frac{P_{sp}^2}{P_0}} \cdot \dfrac{B_o}{B_e} & \text{if } lg = 1 \\[4ex] \dfrac{P_0}{2P_{sp}\frac{\frac{1}{(lg)^k}-1}{1-lg}+\frac{P_{sp}^2}{P_0}\left[\frac{\frac{1}{(lg)^k}-1}{1-lg}\right]^2} \cdot \dfrac{B_o}{B_e} & \text{otherwise} \end{cases}$$

(5.23)

From the above equation, we have the following observations:

1. For $lg < 1$, as $k$ increases, $(\text{ESNR})_{k-g}$ decreases at a dominating rate of $(lg)^{2k}$.

   - $\lim_{k\to\infty}(\text{ESNR})_{k-g} = 0$

2. For $lg = 1$, as $k$ increases, $(\text{ESNR})_{k-g}$ decreases at a rate of $\frac{1}{k^2}$, as shown in Fig. 5-7.

   - $\lim_{k\to\infty}(\text{ESNR})_{k-g} = 0$

3. For $lg > 1$, as $k$ increases, $(\text{ESNR})_{k-g}$ decreases at a rate of $\frac{1}{1-\frac{1}{(lg)^k}}$.

   - $\lim_{k\to\infty}(\text{ESNR})_{k-g} = \dfrac{P_0(lg-1)}{2P_{sp}+\frac{P_{sp}^2}{P_0(lg-1)}} \cdot \dfrac{B_o}{B_e}$

   - If $P_0 \gg P_{sp}$, $\lim_{k\to\infty}(\text{ESNR})_{k-g} = \dfrac{B_o(lg-1)}{2B_e} \cdot \text{OSNR}_{k-g}$

Note that when beat noise of spontaneous emission with itself cannot be ignored, the denominator of ESNR is a quadratic function of the accumulated optical noise

104

power, while the nominator of ESNR is a linear function of the optical signal power, therefore, knowing the ratio of optical signal and optical noise (i.e., the OSNR) is not enough to determine the ESNR. Hence, by knowing the OSNR before the signal is detected is not enough to determine the BER, instead, both exact values of signal power and noise power are needed to derive the BER for both Model k-g and Model k-G.

### 5.5.3 Special case of Model k-g when $lg < 1$ and $lg = \beta$

Define

$$\beta = \frac{P_0}{P_0 + P_{sp}}$$

A special case for $lg < 1$ is when $l_i = \beta/g$ for $i > 1$, and $l_1 = 1/g$. Each fiber span loss is composed of two factors, loss $1/g$ which is compensated by the amplifier gain $g$, and an additional loss $\beta$. Equivalently, we can interpret the loss $\beta$ as an attenuation factor from the gain control of the EDFA, and we arrive at the model in Fig. 5-5.
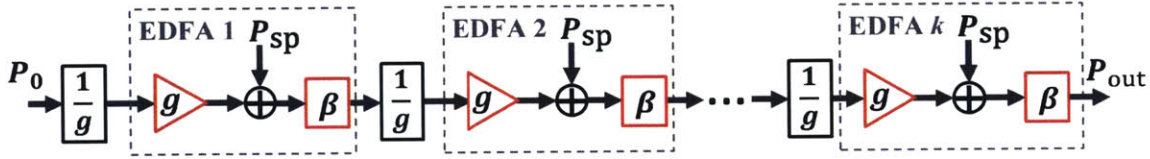


Figure 5-5: A special case of Model k-g, when $lg = \beta$.

Let $(P_{\text{out}})_i$ be the output optical power of the $i$th amplifier in Fig. 5-5. After one amplifier,

$$(P_{\text{out}})_1 = \left(P_0 \frac{1}{g}g + P_{sp}\right)\beta = P_0$$

Therefore, the output power of the first amplifier is controlled to be the same as the original input optical power. Using induction, we can easily prove that

$$(P_{\text{out}})_i = P_0$$

Hence, the effect of having $lg = \beta$ is that the output of each amplifier is controlled to be equal to the initial signal power $P_0$. Note that this is under "constant gain control" that results in equal output power. This is not constant power control, because the output power is controlled to be equal to $P_0$, and if $P_0$ changes, $P_{\text{out}}$ also changes accordingly through $\beta$. If it was under constant power control, $P_{\text{out}}$ would stay the

same despite of variations of $P_0$. Therefore, the OSNR, when the signal is on, is,

$$(\text{OSNR})_{k-g} = \frac{P_0(1-\beta)}{2(g-1)n_{sp}h\nu B_o} \cdot \frac{1}{\left(\frac{1}{\beta}\right)^k - 1}$$

For large $k$, $(\text{OSNR})_{k-g}$ decreases at a rate of $\beta^k$ asymptotically. The ESNR is

$$\text{ESNR}_{k-g} \approx \frac{P_0}{2P_{sp}\frac{\frac{1}{\beta^k}-1}{1-\beta} + \frac{P_{sp}^2}{P_0}\left(\frac{\frac{1}{\beta^k}-1}{1-\beta}\right)^2} \cdot \frac{B_o}{B_e}$$

As $k$ increases, $(\text{ESNR})_{k-g}$ decreases at a rate of $\beta^{2k}$, the same observation as the case of Model k-g with $lg < 1$.

## 5.6 EDFA-Amplified Optical Link Model k-G

We now include the effect of the EDFA random gain. Model k-G in Fig. 5-6 includes intensity noise caused by in-phase gain randomness of the EDFA, and this model uses $G$ instead of $g$.
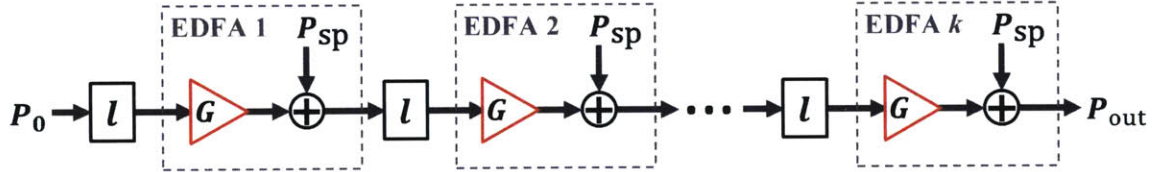


Figure 5-6: Model k-G .

### 5.6.1 Optical Signal and Noise of Model k-G

The average optical signal power after $k$ amplifiers $(P_{\text{sig}})_k$ is

$$(P_{\text{sig}})_k = P_0(lg)^k$$

The average accumulated ASE noise after $k$ amplifiers $(P_{sp})_k$ is

$$(P_{sp})_k = P_{sp}\sum_{i=0}^{k-1}(lg)^i = \begin{cases} kP_{sp} & \text{if } lg = 1 \\ P_{sp}\frac{1-(lg)^k}{1-lg} & \text{otherwise} \end{cases}$$

106

Therefore, the optical signal field after $k$ amplifiers $(E_{\text{sig}})_k$ is

$$(E_{\text{sig}})_k = E_0 \left( \prod_{i=1}^{k} lG_i \right)^{\frac{1}{2}}$$

The accumulated ASE field after $k$ amplifiers $(E_{sp})_k$ is

$$(E_{sp})_k = \sum_{i=1}^{k-1} E_{sp} \left( \prod_{j=i+1}^{k} lG_j \right)^{\frac{1}{2}} + E_{sp}$$

The exact calculation of the intensity noise due to gain randomness depends on the distribution of $G$ and is tedious to calculate. We approximate the intensity noise by two extra variance terms, one from the amplified signal and one from the ASE noise (the derivations are in App. C). The random gain intensity noise from the signal is,

$$\sigma^2_{IG_{sig}} = P_0 (lg)^k \left[ 1 - \left( 1 - \frac{\text{Var}(\sqrt{G})}{g} \right)^k \right]$$

(5.24)

The gain intensity noise from the accumulated ASE is,

$$\sigma^2_{IG_{sp}} = P_{sp} \sum_{i=1}^{k-1} \left( (lg)^{k-i} - \left[ lg \left( 1 - \frac{\text{Var}(\sqrt{G})}{g} \right) \right]^{k-i} \right)$$

$$= \begin{cases} P_{sp} \left\{ (k-1) - \frac{g}{\text{Var}(\sqrt{G})} \left[ 1 - \left( 1 - \frac{\text{Var}(\sqrt{G})}{g} \right)^{k-1} \right] \right\} & \text{if } lg = 1 \\[2ex] P_{sp} \left\{ \frac{lg(1-(lg)^{k-1})}{1-lg} - \frac{lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\left[1 - \left( lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\right)^{k-1}\right]}{1 - lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)} \right\} & \text{if } lg \neq 1, lg \neq \frac{1}{1 - \frac{\text{Var}(\sqrt{G})}{g}} \\[2ex] P_{sp} \left\{ \frac{lg[(lg)^{k-1}-1]}{lg-1} - (k-1) \right\} & \text{if } lg = \frac{1}{1 - \frac{\text{Var}(\sqrt{G})}{g}} \end{cases}$$

(5.25)

Note that with random gain intensity noise considered, the optical field of bit "1" is more random than that of bit "0". We approximate the OSNR when the signal is on by,

$$(\text{OSNR})_{k-G} \triangleq \frac{(P_{\text{sig}})_k}{(P_{sp})_k + \sigma^2_{IG_{sig}} + \sigma^2_{IG_{sp}}}$$

To see how $k$ and random gain affect OSNR and to make the analysis easier, we look at $\frac{1}{\text{OSNR}_{k-G}}$ in three cases where $lg < 1$, $lg = 1$, and $lg > 1$. Notice that $\frac{1}{\text{OSNR}_{k-G}}$

is composed with three parts, and,

$$\frac{1}{\mathrm{OSNR}_{k-G}} = \frac{(P_{sp})_k}{(P_{\mathrm{sig}})_k} + \frac{\sigma^2_{IG_{sig}}}{(P_{\mathrm{sig}})_k} + \frac{\sigma^2_{IG_{sp}}}{(P_{\mathrm{sig}})_k}$$

1. For $lg = 1$,

$$\frac{1}{\mathrm{OSNR}_{k-G}} = \frac{P_{sp}}{P_0}(2k-1) + \left[ \frac{P_{sp}g}{P_0\mathrm{Var}(\sqrt{G})} - \left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right) \right] \left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)^{k-1} + 1 - \frac{P_{sp}g}{P_0\mathrm{Var}(\sqrt{G})}$$
(5.26)

Therefore, for large $k$, $1/\mathrm{OSNR}_{k-G}$ increases at a dominating rate of $2k$ with increasing $k$, or, $\mathrm{OSNR}_{k-G}$ decreases with a dominating rate of $1/2k$ as $k$ increases, and

$$\lim_{k \to \infty} \mathrm{OSNR}_{k-G} = 0$$

2. For $lg < 1$,

$$\frac{1}{\mathrm{OSNR}_{k-G}} = \frac{P_{sp}}{P_0(1-lg)}\left[ \left( \frac{1}{lg} \right)^k - 1 \right] + \left[ 1 - \left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)^k \right] + \frac{\sigma^2_{IG_{sp}}}{(P_{\mathrm{sig}})_k}$$

where

$$\frac{\sigma^2_{IG_{sp}}}{(P_{\mathrm{sig}})_k} = \frac{P_{sp}}{P_0}\left\{ \left[ \frac{lg}{1-lg} - \frac{lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)}{1 - lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)} \right] \left( \frac{1}{lg} \right)^k + \frac{lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)}{1 - lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)} \left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)^{k-1} - \frac{1}{1-lg} \right\}$$

Therefore, for large $k$, $1/\mathrm{OSNR}_{k-G}$ increases at a rate of $\left( \frac{1}{lg} \right)^k$ with increasing $k$, or, $\mathrm{OSNR}_{k-G}$ decreases with a rate of $(lg)^k$ as $k$ increases, and

$$\lim_{k \to \infty} \mathrm{OSNR}_{k-G} = 0$$

3. For $lg > 1$,

$$\frac{1}{\mathrm{OSNR}_{k-G}} = \frac{P_{sp}}{P_0(lg-1)}\left[ 1 - \frac{1}{(lg)^k} \right] + \left[ 1 - \left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)^k \right] + \frac{\sigma^2_{IG_{sp}}}{(P_{\mathrm{sig}})_k}$$

where

$$\frac{\sigma^2_{IG_{sp}}}{(P_{\mathrm{sig}})_k} = \frac{P_{sp}}{P_0}\left\{ \left[ \frac{lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)}{lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right) - 1} - \frac{lg}{lg-1} \right] \frac{1}{(lg)^k} - \frac{lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)}{lg\left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right) - 1} \left( 1 - \frac{\mathrm{Var}(\sqrt{G})}{g} \right)^{k-1} + \frac{lg}{lg-1} \right\}$$

Therefore, for $k$ large, $1/\mathrm{OSNR}_{k-G}$ increases at a dominating rate of

108

$\max\left(a - \left(\frac{1}{lg}\right)^k, b - \left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)^k\right)$ with increasing $k$, where $a$ and $b$ are constants. Or, $\text{OSNR}_{k-G}$ decreases with a dominating rate of $\min\left(\frac{1}{a-\left(\frac{1}{lg}\right)^k}, \frac{1}{b-\left(1-\frac{\text{Var}(\sqrt{G})}{g}\right)^k}\right)$ as $k$ increases. And

$$\lim_{k\to\infty} \text{OSNR}_{k-G} = \frac{P_0(lg - 1)}{2P_{sp}}$$

When considering $\text{OSNR}_{k-G}$, for the cases of both $lg = 1$ and $lg < 1$, the key contributor to signal quality degradation in Model k-G is the accumulated ASE noise, while the effect of the random gain can be neglected; for $lg$ larger than 1, the effect of random gain becomes more significant, but the impact is still non-substantial. With Model k-G when amplifier random gain intensity noise is included, optical bit "1" and bit "0" suffers from different noise power, but the OSNR studied here only considers the noise to bit "1". Therefore, even with coherence detection, because the "0" bit also contributes in determining the threshold, the ratio of the afore-defined OSNR is not enough to calculate the exact BER. Instead, both signal power and noise power for bit "1" and "0" are needed.

### 5.6.2   Electrical Signal and Noise of Model k-G

The photocurrent of bit "1" is,

$$(I_1)_k = \begin{cases} R\left|E_0 + E_{sp}\sqrt{k}\right|^2 + i_{s_1} + i_{\text{sig-sp}_1} + i_{\text{sp-sp}_1} + i_{IG_1} & \text{if } lg = 1 \\ R\left|E_0\sqrt{(lg)^k} + E_{sp}\sqrt{\frac{1-(lg)^k}{1-lg}}\right|^2 + i_{s_1} + i_{\text{sig-sp}_1} + i_{\text{sp-sp}_1} + i_{IG_1} & \text{otherwise} \end{cases}$$

The photocurrent of bit "0" is,

$$(I_0)_k = \begin{cases} Rk|E_{sp}|^2 + i_{s_0} + i_{\text{sp-sp}_0} + i_{IG_0} & \text{if } lg = 1 \\ R\frac{1-(lg)^k}{1-lg}|E_{sp}|^2 + i_{s_0} + i_{\text{sp-sp}_0} + i_{IG_0} & \text{otherwise} \end{cases}$$

Therefore, for bit "1", we have,

$$(\bar{I}_1)_k = \begin{cases} R(P_0 + kP_{sp}) & \text{if } lg = 1 \\ R\left[P_0(lg)^k + P_{sp}\frac{1-(lg)^k}{1-lg}\right] & \text{otherwise} \end{cases}$$

$$(\sigma_{s_1}^2)_k = \begin{cases} 2qR(P_0 + kP_{sp})B_e & \text{if } lg = 1 \\ 2qR\left[P_0(lg)^k + P_{sp}\frac{1-(lg)^k}{1-lg}\right]B_e & \text{otherwise} \end{cases}$$

$$(\sigma_{sig-sp_1}^2)_k = \begin{cases} 4R^2 P_0 k(g-1)n_{sp}h\nu B_e & \text{if } lg = 1 \\ 4R^2 P_0(g-1)n_{sp}h\nu B_e\frac{(lg)^k\left[1-(lg)^k\right]}{1-lg} & \text{otherwise} \end{cases} \qquad (5.27)$$

$$(\sigma_{sp-sp_1}^2)_k = \begin{cases} 4R^2\left[k(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{if } lg = 1 \\ 4R^2\left[\frac{1-(lg)^k}{1-lg}(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{otherwise} \end{cases} \qquad (5.28)$$

$$(\sigma_{IG_1}^2)_k = \sigma_{IG_{sig-sig}}^2 + \sigma_{IG_{sig-sp}}^2 + \sigma_{IG_{sp-sp}}^2$$

where $\sigma_{IG_{sig-sig}}^2$ is the random gain noise from signal power, $\sigma_{IG_{sig-sp}}^2$ is the random gain noise from beat noise of signal with ASE, and $\sigma_{IG_{sp-sp}}^2$ is the random gain noise from the beat noise of ASE with itself.

For bit "0", we have

$$(\bar{I}_0)_k = \begin{cases} kRP_{sp} & \text{if } lg = 1 \\ RP_{sp}\frac{1-(lg)^k}{1-lg} & \text{otherwise} \end{cases}$$

$$(\sigma_{s_0}^2)_k = \begin{cases} 2kqRP_{sp}B_e & \text{if } lg = 1 \\ 2qRP_{sp}B_e\frac{1-(lg)^k}{1-lg} & \text{otherwise} \end{cases}$$

$$(\sigma_{sp-sp_0}^2)_k = \begin{cases} 4R^2\left[k(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{if } lg = 1 \\ 4R^2\left[\frac{1-(lg)^k}{1-lg}(g-1)n_{sp}h\nu\right]^2 B_o B_e & \text{otherwise} \end{cases}$$

$$(\sigma_{IG_0}^2)_k = \sigma_{IG_{sp-sp}}^2$$

The random gain intensity noise is derived in App. C, and,

$$\sigma_{IG_{sig-sig}}^2 = R^2 P_0^2 (lg)^{2k}\left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^k - 1\right] \qquad (5.29)$$

110

$$\sigma^2_{IG_{sig-sp}} \geq 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot \sum_{j=0}^{k-1} \left\{ \left[ lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) \right]^j - (lg)^j \right\}$$

$$= \begin{cases} 4R^2 P_0 S_{sp} B_e \cdot \left\{ \frac{g^2}{\sigma_G^2} \left[ \left( 1 + \frac{\sigma_G^2}{g^2} \right)^k - 1 \right] - k \right\} & \text{if } lg = 1 \qquad (5.30) \\[3mm] 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot \left\{ \frac{\left[ lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) \right]^k - 1}{lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) - 1} - \frac{(lg)^k - 1}{lg - 1} \right\} & \text{otherwise} \end{cases}$$

If $lg \neq 1$,

$$\sigma^2_{IG_{sp-sp}} > 4R^2 S_{sp}^2 B_o B_e \cdot \sum_{i=1}^{k-1} \left\{ \frac{\left[ (lg)^2 \left( 1 + \frac{\sigma_G^2}{g^2} \right) \right]^{k-i}}{lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) - 1} - \frac{(lg)^{2(k-i)}}{lg - 1} + \frac{lg \frac{\sigma_G^2}{g^2} (lg)^{k-i}}{\left[ lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) - 1 \right](lg - 1)} \right\}$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot \left\{ \frac{(lg)^2 \left( 1 + \frac{\sigma_G^2}{g^2} \right)}{lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) - 1} \cdot \frac{\left[ (lg)^2 \left( 1 + \frac{\sigma_G^2}{g^2} \right) \right]^{k-1} - 1}{(lg)^2 \left( 1 + \frac{\sigma_G^2}{g^2} \right) - 1} - \frac{(lg)^2}{lg - 1} \cdot \frac{(lg)^{2(k-1)} - 1}{(lg)^2 - 1} + \frac{(lg)^2 \frac{\sigma_G^2}{g^2} \left[ (lg)^{k-1} - 1 \right]}{\left[ lg \left( 1 + \frac{\sigma_G^2}{g^2} \right) - 1 \right](lg - 1)^2} \right\}$$
$$(5.31)$$

If $lg = 1$,

$$\sigma^2_{IG_{sp-sp}} = 4R^2 S_{sp}^2 B_o B_e \cdot \left\{ \left( \frac{g^2}{\sigma_G^2} \right)^2 \left[ \left( 1 + \frac{\sigma_G^2}{g^2} \right)^k - \left( 1 + \frac{\sigma_G^2}{g^2} \right) \right] - \frac{g^2}{\sigma_G^2}(k-1) - \frac{k(k-1)}{2} \right\}$$
$$(5.32)$$

Modifying the ESNR defined in Eq. (5.18) to include the random gain intensity noise in Model k-G as (shot noise is ignored),

$$(\text{ESNR})_{k-G} \triangleq \frac{\left( (\bar{I}_1)_k - (\bar{I}_0)_k \right)^2}{\left( \sigma^2_{sig-sp_1} \right)_k + \left( \sigma^2_{sp-sp_1} \right)_k + (\sigma^2_{IG_1})_k}$$

To make the analysis more straightforward, let's look into $1/(\text{ESNR})_{k-G}$ instead, and,

$$\frac{1}{(\text{ESNR})_{k-G}} = \frac{1}{(\text{ESNR})_{k-g}} + \frac{(\sigma^2_{IG_1})_k}{R^2 P_0^2 (lg)^{2k}}$$

where $(\text{ESNR})_{k-g}$ is defined in Eq. (5.22). Note that

$$\frac{(\sigma^2_{IG_1})_k}{R^2 P_0^2 (lg)^{2k}} = \frac{\sigma^2_{IG_{sig-sig}}}{R^2 P_0^2 (lg)^{2k}} + \frac{\sigma^2_{IG_{sig-sp}}}{R^2 P_0^2 (lg)^{2k}} + \frac{\sigma^2_{IG_{sp-sp}}}{R^2 P_0^2 (lg)^{2k}}$$

From Eq. (5.29), we have,

$$\frac{\sigma^2_{IG_{sig-sig}}}{R^2 P_0^2 (lg)^{2k}} = \left( 1 + \frac{\sigma_G^2}{g^2} \right)^k - 1 \qquad (5.33)$$

111

From Eq. (5.30), we have,

$$
\frac{\sigma^2_{IG_{sig-sp}}}{R^2 P_0^2 (lg)^{2k}} \geq
\begin{cases}
\frac{4 S_{sp} B_e}{P_0} \cdot \left\{ \frac{g^2}{\sigma_G^2} \left[ \left(1 + \frac{\sigma_G^2}{g^2}\right)^k - 1 \right] - k \right\} & \text{if } lg = 1 \\[4ex]
\frac{4 S_{sp} B_e}{P_0} \cdot \left\{ \frac{\left[ \left(1 + \frac{\sigma_G^2}{g^2}\right) \right]^k - \frac{1}{(lg)^k}}{lg \left(1 + \frac{\sigma_G^2}{g^2}\right) - 1} - \frac{1 - \frac{1}{(lg)^k}}{lg - 1} \right\} & \text{otherwise}
\end{cases}
\tag{5.34}
$$

From Eq. (5.31) and (5.32), we have,

$$
\frac{\sigma^2_{IG_{sp-sp}}}{R^2 P_0^2 (lg)^{2k}} \geq
\begin{cases}
\frac{4 S_{sp}^2 B_o B_e}{P_0^2} \cdot \left\{ \left(\frac{g^2}{\sigma_G^2}\right)^2 \left[ \left(1 + \frac{\sigma_G^2}{g^2}\right)^k - \left(1 + \frac{\sigma_G^2}{g^2}\right) \right] - \frac{g^2}{\sigma_G^2}(k-1) - \frac{k(k-1)}{2} \right\} & \text{if } lg = 1 \\[4ex]
\frac{4 S_{sp}^2 B_o B_e}{P_0^2} \cdot \left\{ \frac{(lg)^2 \left(1 + \frac{\sigma_G^2}{g^2}\right)}{lg\left(1+\frac{\sigma_G^2}{g^2}\right)-1} \cdot \frac{\frac{1}{(lg)^2}\left[\left(1+\frac{\sigma_G^2}{g^2}\right)\right]^{k-1} - \frac{1}{(lg)^{2k}}}{(lg)^2\left(1+\frac{\sigma_G^2}{g^2}\right)-1} - \frac{1 - \frac{1}{(lg)^{2(k-1)}}}{(lg-1)\left[(lg)^2-1\right]} + \frac{(lg)^2 \frac{\sigma_G^2}{g^2}\left[\frac{1}{(lg)^{k+1}} - \frac{1}{(lg)^{2k}}\right]}{\left[lg\left(1+\frac{\sigma_G^2}{g^2}\right)-1\right](lg-1)^2} \right\} & \text{o.w.}
\end{cases}
$$

(5.35)

From the above equations, we have the following observations:

1. For $lg < 1$, as $k$ increases, $1/(\text{ESNR})_{k-G}$ is dominated by the term $\max\left( \frac{1}{\text{ESNR}_{k-g}}, \frac{\sigma^2_{IG_{sig-sig}}}{R^2 P_0^2 (lg)^{2k}} \right)$, which can be re-expressed as, $\max\left( a\left(\frac{1}{lg}\right)^k, b\left(\frac{1}{lg}\right)^{2k}, c\left(1 + \frac{\sigma_G^2}{g^2}\right)^k \right)$, where $a, b$, and $c$ are constants w.r.t. $k$ and $\frac{\sigma_G^2}{g^2}$.

   - $\lim_{k\to\infty}(\text{ESNR})_{k-G} = 0$

2. For $lg = 1$, as $k$ increases, $1/(\text{ESNR})_{k-G}$ increases at a dominating rate of $\left(1 + \frac{4 S_{sp} B_e}{P_0} \cdot \frac{g^2}{\sigma_G^2} + \frac{4 S_{sp}^2 B_o B_e}{P_0^2} \cdot \left(\frac{g^2}{\sigma_G^2}\right)^2 \right)\left(1 + \frac{\sigma_G^2}{g^2}\right)^k$, as shown in Fig. 5-7.

   - $\lim_{k\to\infty}(\text{ESNR})_{k-G} = 0$

3. For $lg > 1$, as $k$ increases, $1/(\text{ESNR})_{k-G}$ increases at a rate of $\left(1 + \frac{4 S_{sp} B_e}{P_0} \cdot \frac{1}{lg\left(1+\frac{\sigma_G^2}{g^2}\right)-1} + \frac{4 S_{sp}^2 B_o B_e}{P_0^2} \cdot \frac{\left(1+\frac{\sigma_G^2}{g^2}\right)}{\left[lg\left(1+\frac{\sigma_G^2}{g^2}\right)-1\right]\left[(lg)^2\left(1+\frac{\sigma_G^2}{g^2}\right)-1\right]} \right)\left(1 + \frac{\sigma_G^2}{g^2}\right)^k$.

   - $\lim_{k\to\infty}(\text{ESNR})_{k-G} = 0$

Note that when amplifier random gain noise cannot be ignored, the denominator of ESNR is a sum of a quadratic function of the accumulated optical noise power and noise components from the random gain noise contribution. Therefore, even if the exact values of both optical signal and total optical noise are known, they are not enough to derive the BER, because the optical noise includes both spontaneous emission noise

112

and the random gain noise and the electrical noise is not a simple linear function of the total optical noise. Therefore, the exact composition of optical noise is needed to derive the BER.
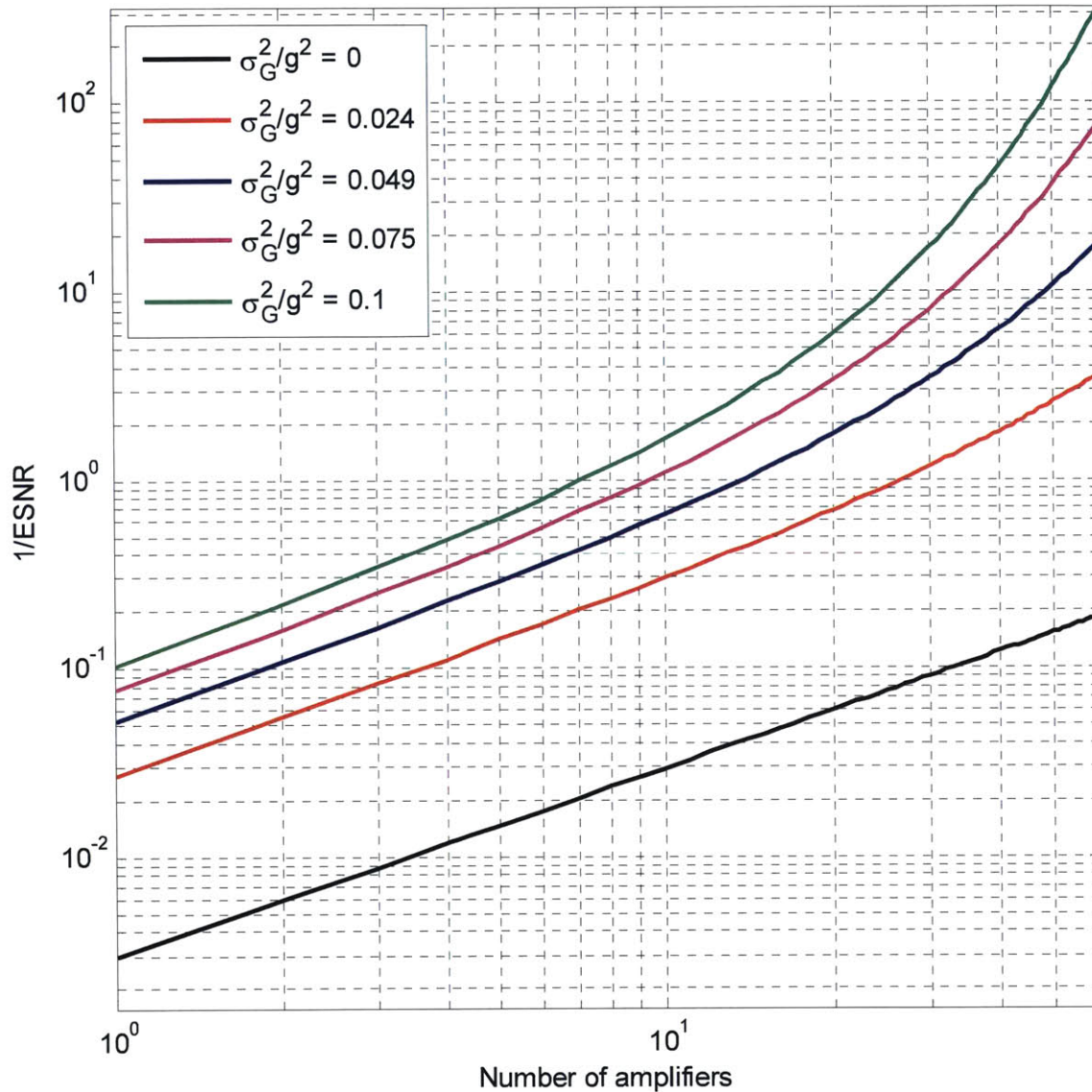
Figure 5-7: The inverse of ESNR for Model k-G and Model k-g, in log-log scale. Note that $\frac{1}{\text{ESNR}_{k-g}}$ is plotted by the black curve with $\sigma_G^2/g^2 = 0$. Parameters used in this plot are: $lg = 1$, signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.

## 5.7 Numerical Examples of Model k-g and Model k-G

As discussed in the previous section, with presence of random gain intensity noise, OSNR or ESNR is not enough to determine the BER. Next we look into the dependence of BER on the effect of a chain of amplifiers and the random gain intensity noise using numerical examples for the case of matched loss and gain where $lg = 1$. In the exapmles, we considered a chain of 60 amplifiers, with parameters: initial signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.

Figure 5-8 and Fig. 5-9 plot the BER as a function of the number of amplifiers in normal scale and log-log scale, respectively. Clearly, with the presence of random gain intensity noise, BER is higher than when there is no random gain noise. In addition, BER increases faster as a function of the number of amplifiers when the random gain noise is larger. Figure 5-10 plots the BER ratio of Model k-G and Model k-g as a function of the number of amplifiers. It shows the ratio $\frac{\text{BER}_{k-G}}{\text{BER}_{k-g}}$ decreases with increase of number of amplifiers, indicating that the effect of random gain noise is more significant with a link of less amplifiers. This is because, although BER increases as noise increases, the rate of increase of the BER decreases with increase of noise, and the BER approaches the horizontal asymptote 0.5 as noise increases. When there are more amplifiers, the BER is already very high due to accumulation of noise, and the increase of the BER becomes slower, therefore, the presence of random gain noise leads to a smaller $\frac{\text{BER}_{k-G}}{\text{BER}_{k-g}}$ with more amplifiers.

Figure 5-11 plots the BER of Model k-G as a function of the ratio of the variance and mean-squared of the amplifier gain (i.e., $\sigma_G^2/g^2$). Fig. 5-12 plots the ratio of BERs of Model k-G and Model k-g (i.e., $\frac{\text{BER}_{k-G}}{\text{BER}_{k-g}}$) as a function of $\sigma_G^2/g^2$. Both BER and $\frac{\text{BER}_{k-G}}{\text{BER}_{k-g}}$ increases with more randomness of the gain. $\frac{\text{BER}_{k-G}}{\text{BER}_{k-g}}$ is larger with less number of amplifiers.
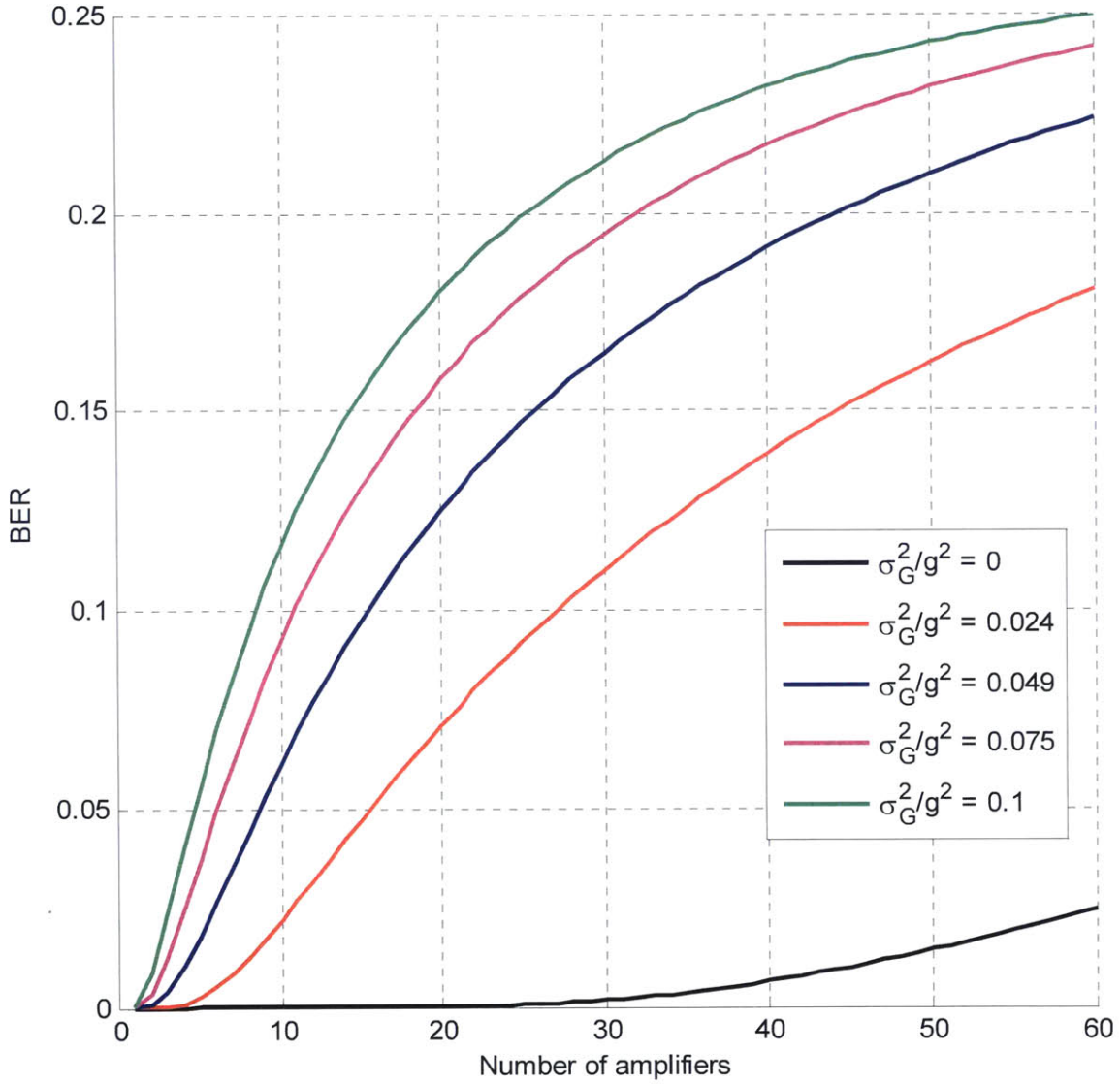
Figure 5-8: BER of Model k-G (i.e., $\mathrm{BER}_{k-G}$) as a function of the number of amplifiers, in normal scale. Note the black curve with $\sigma_G^2/g^2 = 0$ corresponds to the BER of Model k-g. Parameters used in this plot are: $lg = 1$, signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.

116
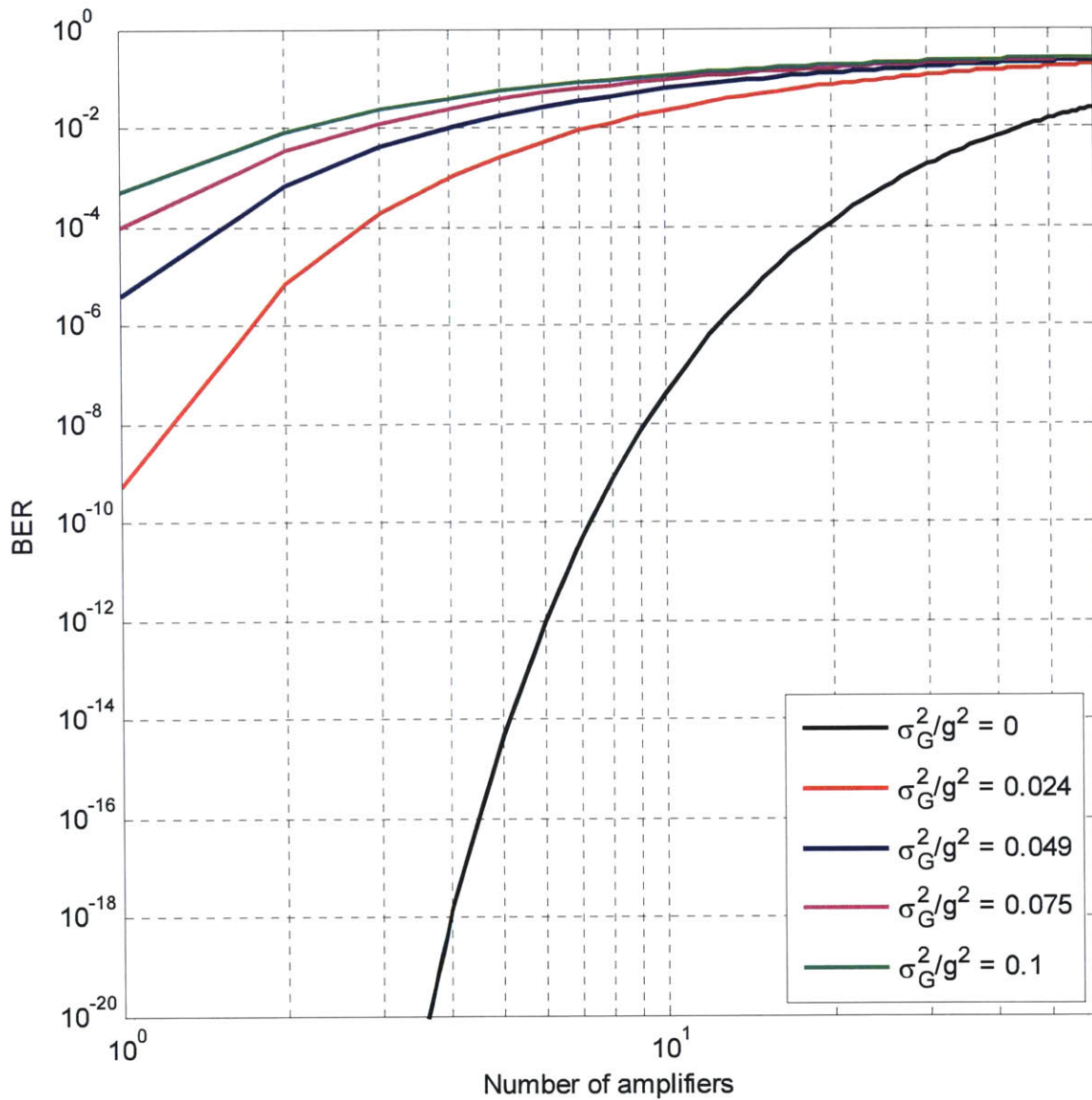
Figure 5-9: BER of Model k-G as a function of number of amplifiers, in log-log scale. Note the black curve with $\sigma_G^2/g^2 = 0$ corresponds to the BER of Model k-g. Parameters used in this plot are: $lg = 1$, signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.
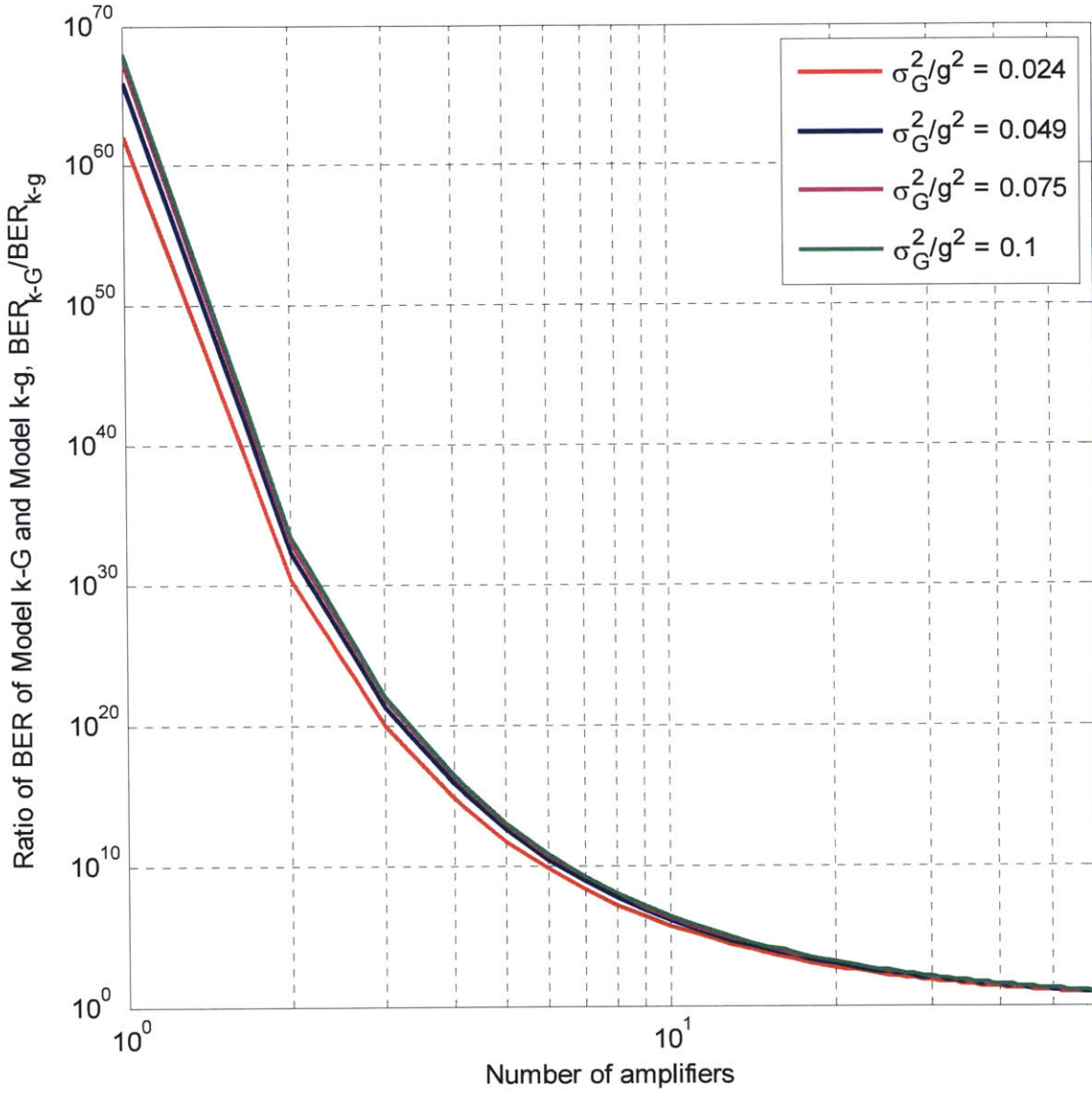
Figure 5-10: Ratio of BERs of Model k-G and Model k-g (i.e., $\frac{\text{BER}_{k-G}}{\text{BER}_{k-g}}$) as a function of number of amplifiers, in log-log scale. Parameters used in this plot are: $lg = 1$, signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.
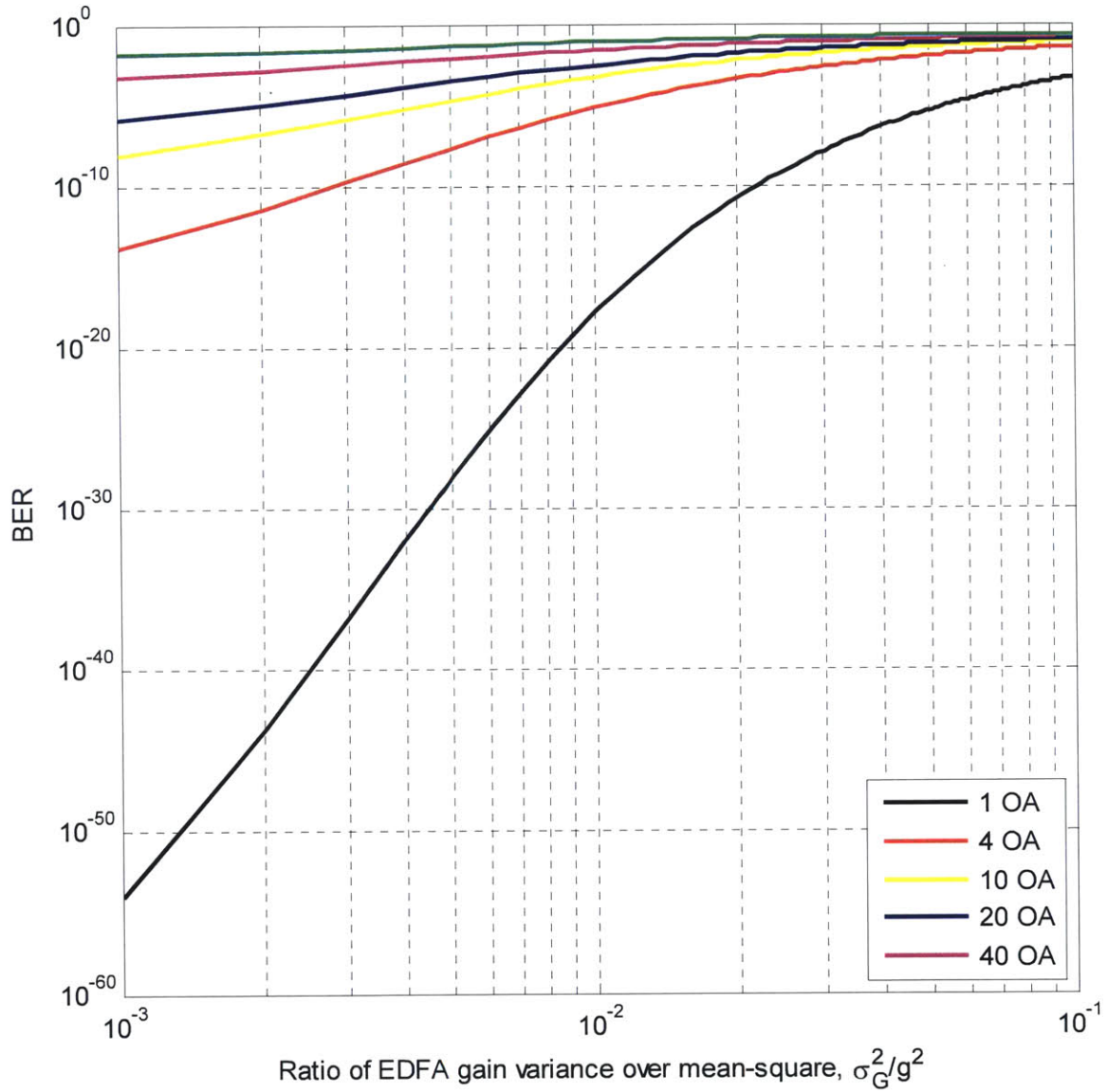
Figure 5-11: BER of Model k-G as a function of the ratio of the variance and mean-squared of the amplifier gain (i.e., $\sigma_G^2/g^2$), in log-log scale. Parameters used in this plot are: $lg = 1$, signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.
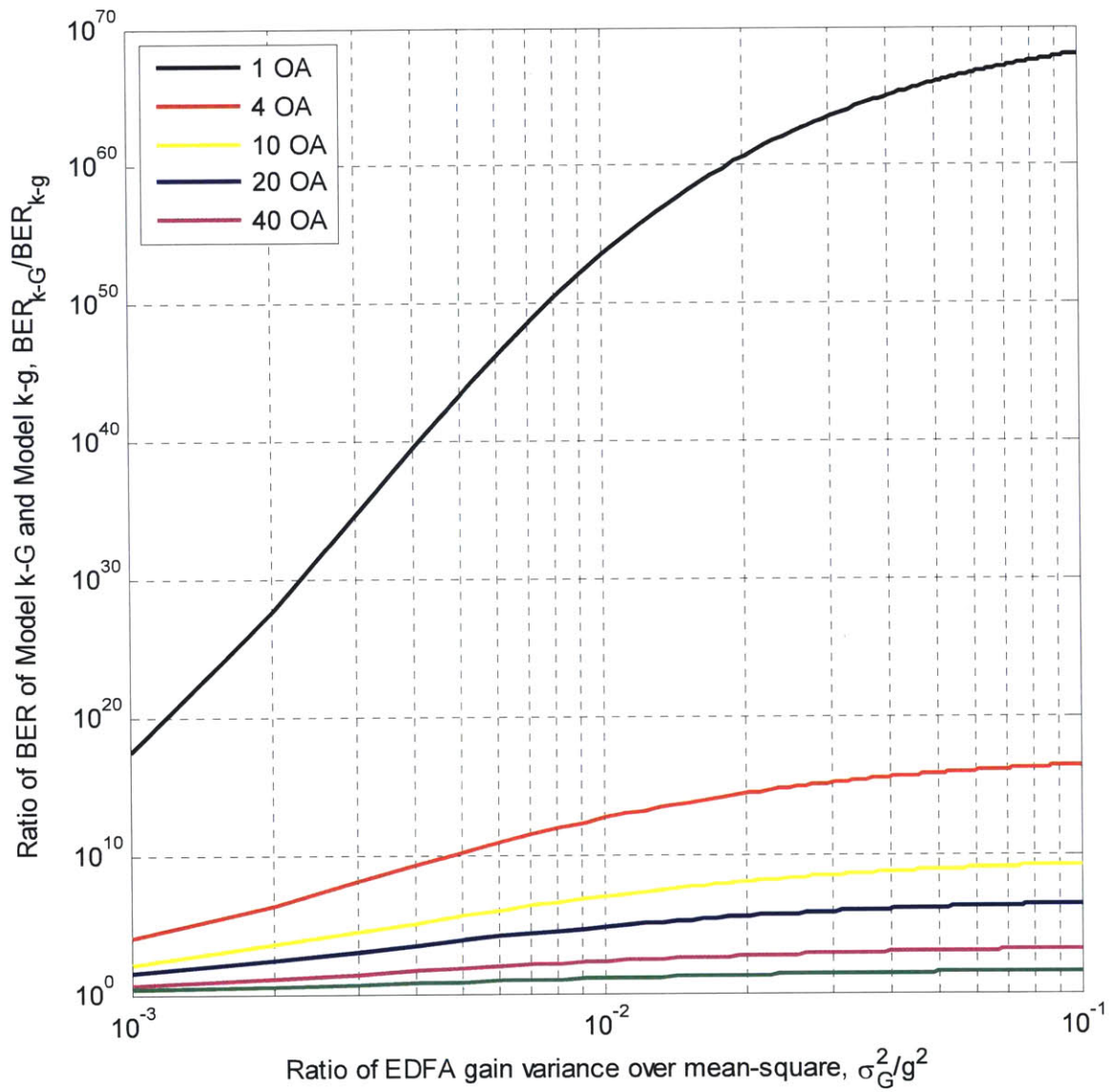
Figure 5-12: Ratio of BER of Model k-G and BER of Model k-g (i.e., $\frac{\mathrm{BER}_{k-G}}{\mathrm{BER}_{k-g}}$) as a function of the variance and mean-squared of the amplifier gain (i.e., $\frac{\sigma_G^2}{g^2}$), in log-log scale. Parameters used in this plot are: $lg = 1$, signal power $P_0 = 5$ mW, ASE noise of one amplifier $P_{sp} = 0.01$ mW, detector optical bandwidth $B_o = 15$ GHz, detector electrical bandwidth $B_e = 11$ GHz, and detector responsivity $R = 3.73$ C/J assuming 1542 nm wavelength and 100% efficiency with $\eta = 1$.

120

## 5.8 Summary and Discussions of Experimental Results

The experiment discussed in this section was carried out by Joseph Junio at Bell Lab with partial experiment results published in [23].

### 5.8.1 Experiment Setup

As shown in Fig. 5-13, a testbed consisting of a line network with optical switches at intermediate nodes was set up to represent a segment (an end-to-end path composed of three links) of a large-scale mesh network. Wavelength selective switches (WSS) represent network nodes, with WSS 1 for the source node, and WSS 2 and WSS 3 for intermediate nodes. At each WSS node, attenuators are controlled to ensure equal power across the spectrum is launched into the next link. Each optical link is amplified using three EDFAs, which are separated by a 40 km span of standard single mode fiber for ~21 dB loss per span. Two-stage EDFAs with mid-stage variable optical attenuators and "constant gain" controllers were used and were capable of adjusting the internal gain to within 0.1 dB of the target within milliseconds. Switch 1 can be turned on/off to simulate the on and off of the switched channels. It is driven by a smooth voltage function from a programmable waveform function to control the switching on/off functions of lightpaths.
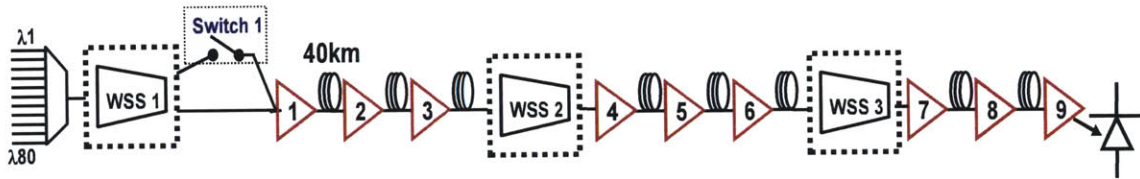


Figure 5-13: Experimental setup depicting WSS Nodes 1, 2, 3 and EDFAs 1-9, separated by 40 km spans of SSMF. EDFA 9 output is amplified by EDFA 10 and filtered by WSS 4 (both not shown).

### 5.8.2 Switching-Induced Fast Transients

Two types of switching, step-function switching and adiabatic switching (Fig. 5-14), were used. For the adiabatic switching, voltage of a raised cosine function was used to drive the switch, and impacts of different switching durations (from 10 ns to 40 ms) were investigated.

121
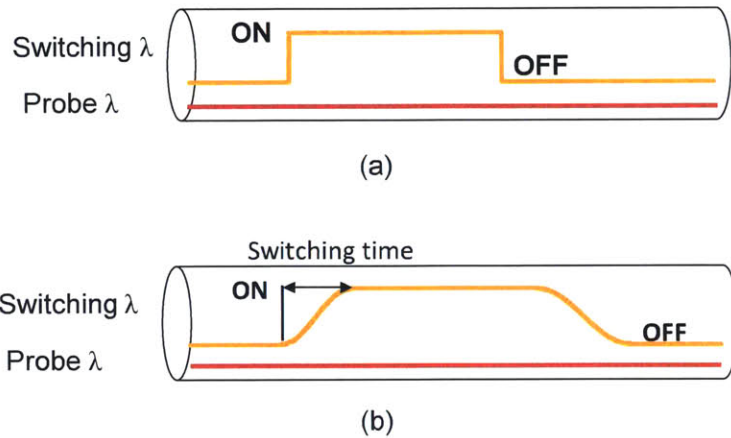
Figure 5-14: (a) Step switching function. (b) Adiabatic switching function of a raised-cosine function.

## Step-Function Switching

When the switching channel is switched on/off abruptly, there is a sudden short-age/excess of upper stage population in the EDFA, and the probe channel output power experiences an abrupt down-shoot/overshoot. This happens faster than the response time of EDFA control. Later (after ~5 ms) the "constant gain control" of EDFA kicks in and brings the gain back to the pre-set value. In Fig. 5-15, a transient of height > 5 dB lasting ~10 ms was observed.
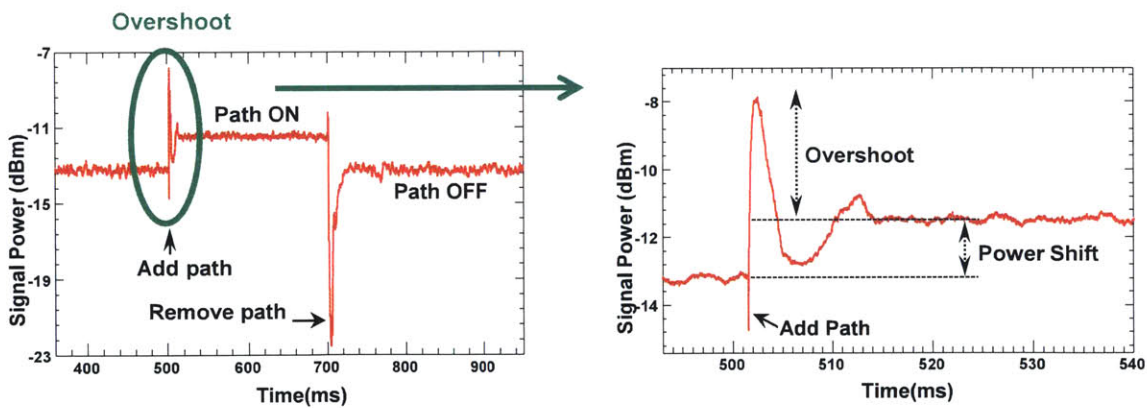


Figure 5-15: (Left) A probe channel output after 9 EDFAs when there are 4 channels present and 4 channels are added and then dropped that corresponds to the worst case channel configurations of adding/dropping 4 channels. (Right) The initial turn-on transient with expanded time scale showing the transients subsides after ~5 ms.

122

## Adiabatic Switching

Figure 5-16 and Fig. 5-17 show the transient duration over a chain of many EDFAs (10) indicating switching times of ~5 ms are long enough to quench transients without detailed sensing of the state of the EDFA physical dynamics and control systems. Since the optical flows in OFS persist for one second or more, we can afford several milliseconds to switch on/off lightpaths without significantly sacrificing throughput performance. By switching wavelengths adiabatically (slowly) within 5 ms, the amplifier dynamics and the EDFA control systems can respond smoothly, reducing the severity of the optical power transients. The adiabatic switching should quench most of the transients except for rare cases such as multiple (> 2) switch-on/offs occurring at the same time. A Transport Layer protocol [21] has been designed and will correct the resulting errors in transmission.
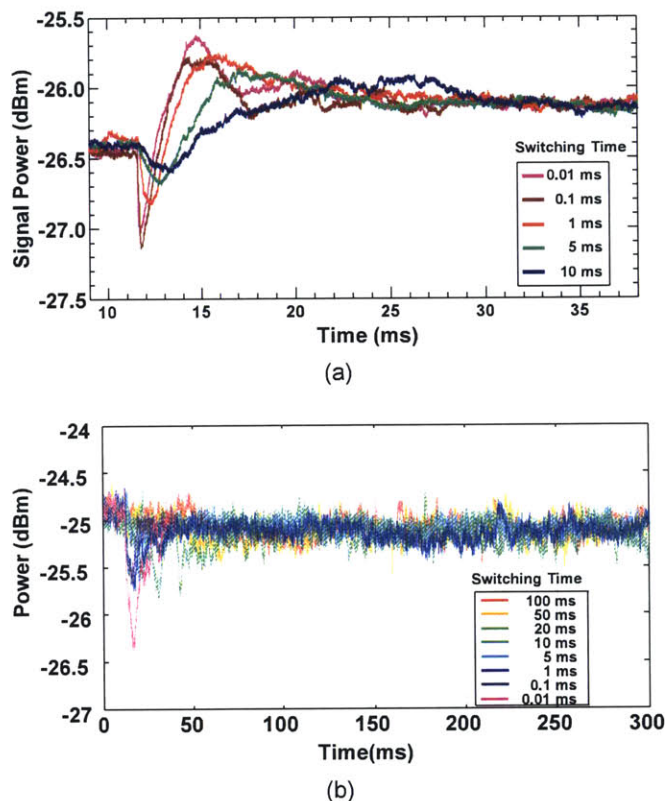


Figure 5-16: (a) Transient events on probe channel (1561.01 nm) when 2 channels are present and one channel added for various turn-on times. (b) Transient events on probe channel (1561.01 nm) when 4 channels are present and one channel dropped for various turn-off times.

123
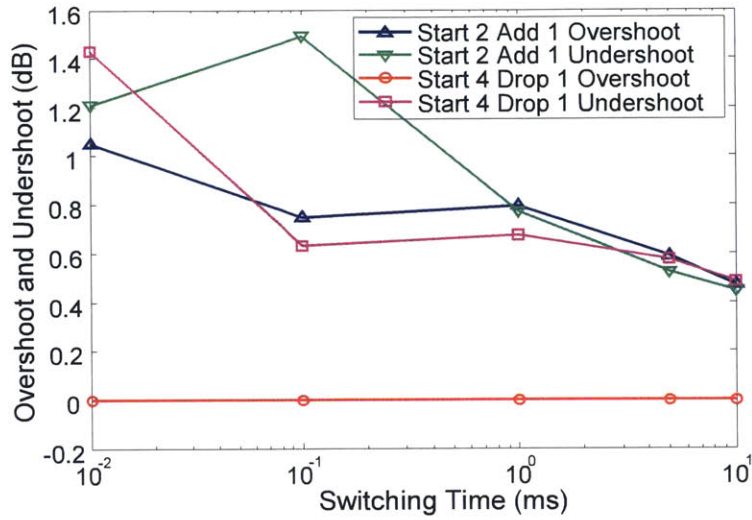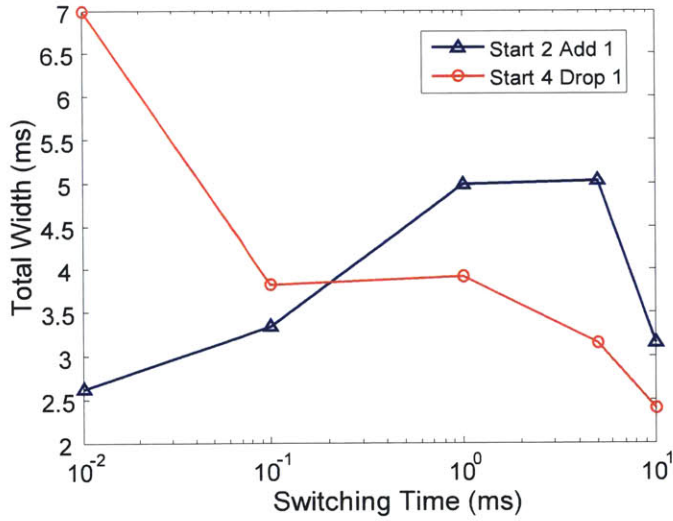
Figure 5-17: (a) Peak value of transients and (b) 3 dB width of both turn-on and turn-off transients versus switching on/off times.

### 5.8.3 Channel Quality Degradations

A 10Gbps On-Off Keying signal with extinction ratio 16:1 in dB was used in the experiment. Channel quality and BER were investigated w.t.r. stages of amplifications (number of amplifiers) and channel configuration (number of channels that are switched On in the fiber).

Light passing through long-haul fiber experiences noise from numerous sources, and

both received "1" and "0" bit can be assumed to have a Gaussian distribution.

Figure 5-18 shows the eye pattern of the communication link at the output of the detector. There is a significant closure of the eye due to the noise accumulated over the link of 9 amplifiers. Figure 5-19 shows the Gaussian statistics of the "1" and "0" bits. The "0" bit has non-zero mean because of both the non-zero extinction ratio and the accumulated ASE noise. In addition, the variance for the "1" bit is larger than that of the "0" bit.
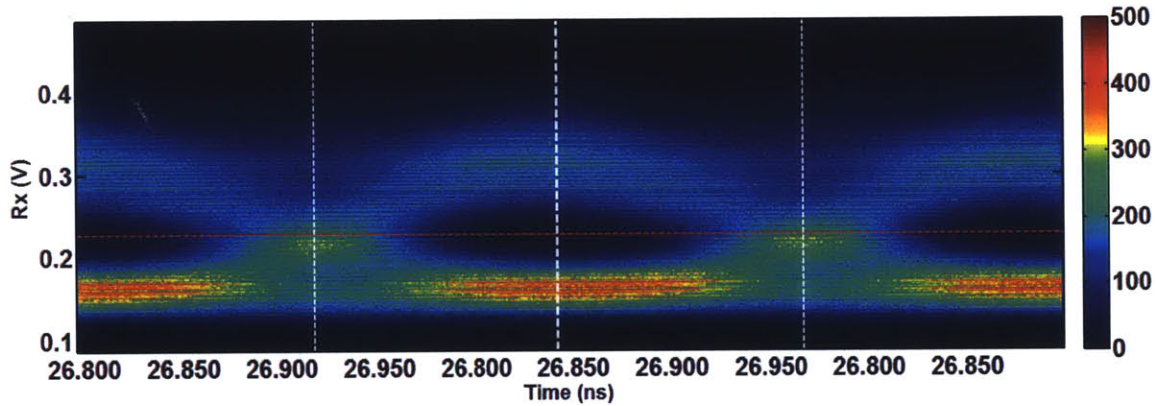
Figure 5-18: Eye pattern of communication link after a chain of 9 EDFAs, showing significant eye closure.
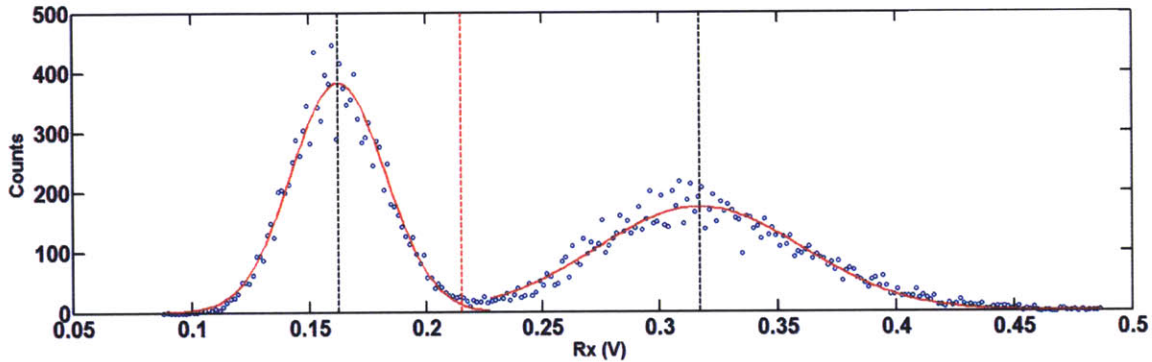
Figure 5-19: Sampled statistics of bit "1" and "0" after a 9-EDFA chain. The mean for bit "0" is non-zero and the variance for bit "0" is less than that for bit "1".

## Signal Statistics after a Cascade of Amplifiers

A standard BER tester (Anritsu) was used to measure the error probability in terms of the bit error rate for a $2^{15} - 1$, 10 Gbits/s OOK PRBS bit stream at the link egress. A photo-detector was used to measure the statistics for the "1" and "0" levels as inputs

125

to the model. Figure 5-20 shows the variance/mean-squared as a function of number of amplifiers that signal goes through for the "1" bit (signal ON), with channel load of 2, 6, 10, and 20 channels. Figure 5-21 shows the variance/mean-squared as a function of channel load for the "1" bit at different locations along the path (after 2, 4, 7, and 10 amplifiers, respectively). Fig. 5-22 shows the variance/mean-squared of the "1" and "0" levels of the signal channel (measured after 10 amplifiers) as more channels are added into the link. The data in both Fig. 5-20 and 5-21 shows two important observations:

1. Variance/mean-squared decreases with increasing channel load, indicating improving channel performance.

2. Variance/mean-squared increases approximately exponentially with number of amplifiers, indicating channel quality degradation with more amplifiers.

The bit error rate, BER (Fig. 5-23), as expected, improves as the number of wavelength channels present increases. Figure 5-24 plotted BERs predicted by the single-threshold detection model in Section 5.1 against the BER measured directly, indicating the single-threshold detector is a good approximation of the optimal detector.
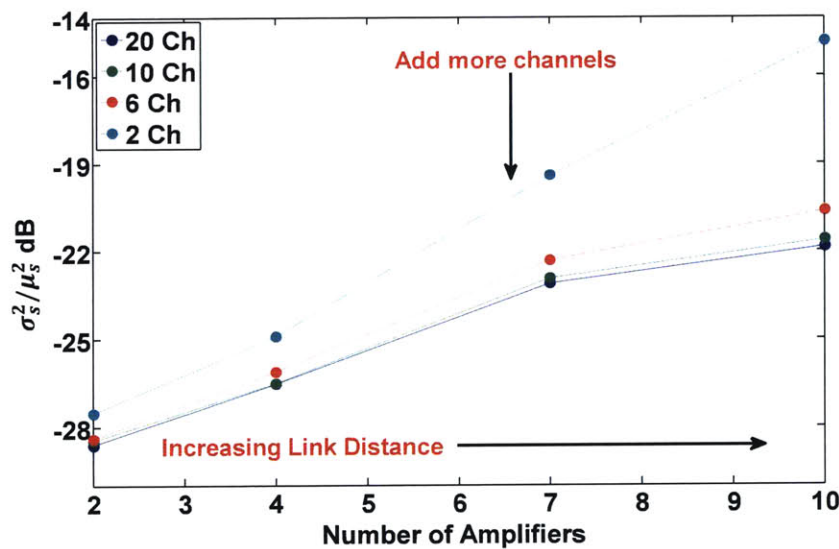


Figure 5-20: Variance/mean-squared of the "1" bit as a function of number of amplifiers, at each channel configuration. The slopes increase for more amplifiers, but decrease as more channels are added.

126
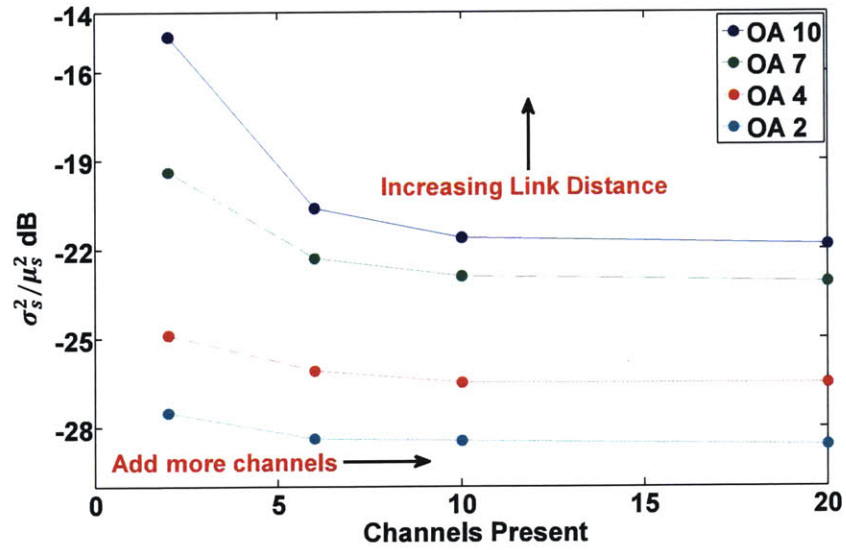
Figure 5-21: Variance/mean-squared of the "1" bit as a function of channel configuration, at each amplifier.
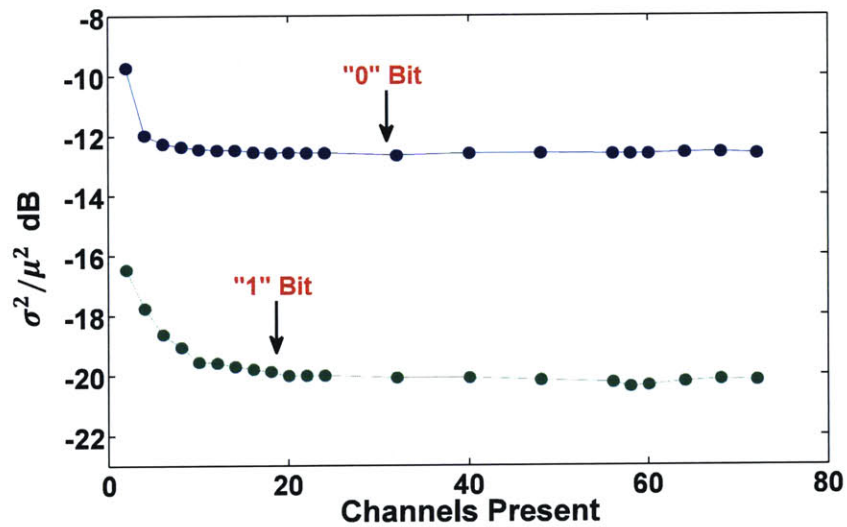


Figure 5-22: Variance/mean squared of the "1" bit and "0" bit measured after 9 amplifiers in the link, plotted in dB scale.
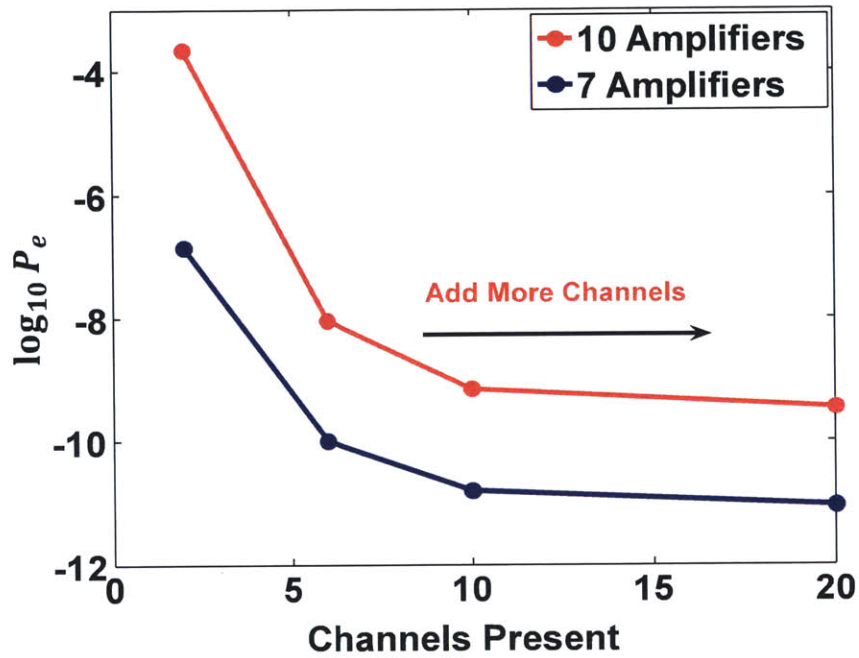
Figure 5-23: Error probability computed from the variances in Fig. 5-20 at the output of amplifiers 7 and 10.
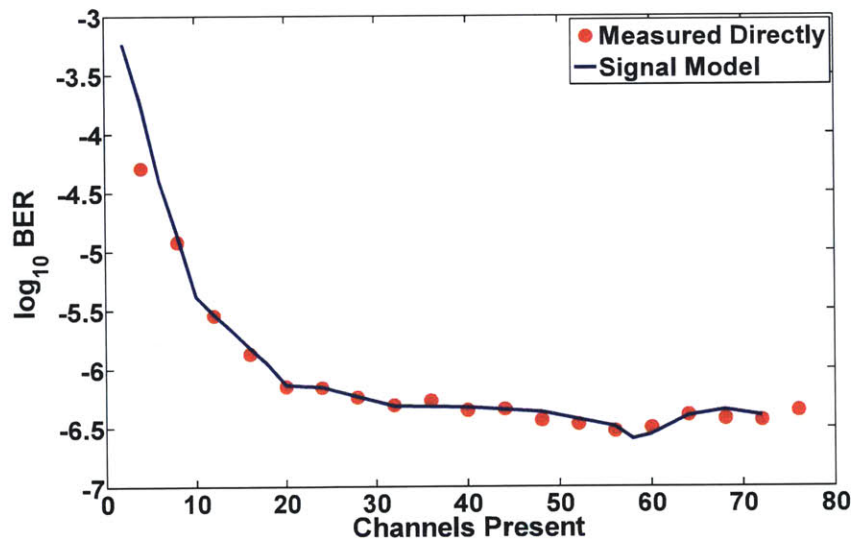


Figure 5-24: Bit error probability, BER as a function of channels present in the link. Points are experimentally measured (10 Gbits/s, OOK), solid line are calculated using the single-threshold detection model in Section 5.1 and the measured photocurrents for the one and zero bit levels.

One possible explanation of Obaservation 1 is from Eq. (5.11), and $N_1 + N_2 = N_t$ where $N_t$ is the total ion density. With these two equations, the ASE noise can be

128

written as,

$$P_{sp} = 2h\nu B_o(g-1)\left(\frac{1}{2} + \frac{\frac{1}{2}N_t}{2N_2 - N_t}\right)$$

When there are more channels present in the fiber, to maintain the same gain $g$, the controller cause the pump to raise the inversion level, promoting more carriers to the upper state. In addition, the increase in signal from the added wavelengths leads to an increase of stimulated emission, draining the gain and reducing spontaneously emitted photons. Thus, as the number of wavelengths in the fiber increases, the spontaneous emission noise in the fiber decreases. From Eq. (5.20), (5.21), (5.27), (5.28), (5.30), and (5.31), for both Model k-g and Model k-G, there are both linear terms of $P_{sp}$ and quadratic terms of $P_{sp}$ in the variance of bit "1". Therefore, $\sigma_s^2/\mu_s^2$ is a quadratic function of $P_{sp}$. Assume there is a linear relationship between the upper state carrier population of the EDFA $N_2$ and the number of "On" channels in a fiber $N_{ch}$, then $\sigma_s^2/\mu_s^2$ can be written as,

$$\frac{\sigma_s^2}{\mu_s^2} = \frac{a}{N_{ch} - b} + \frac{c}{(N_{ch} - b)^2} + d \qquad (5.36)$$

where $a$, $b$, $c$ and $d$ are fitting parameters. Parameter $b$ is from $N_t$, and parameters $a$, $c$ and $d$ account for the effects of $P_{sp}$, the EDFA random gain, and the fiber loss, respectively. Figure 5-25 plots $\sigma_s^2/\mu_s^2$ as a function of the number of amplifiers using both the experimental data and the matched data using minimum mean-squared estimated parameters $a$, $b$, $c$ and $d$, showing a good agreement between these two data sets. Therefore, our conjecture offers a possible explanation of Observation 1.
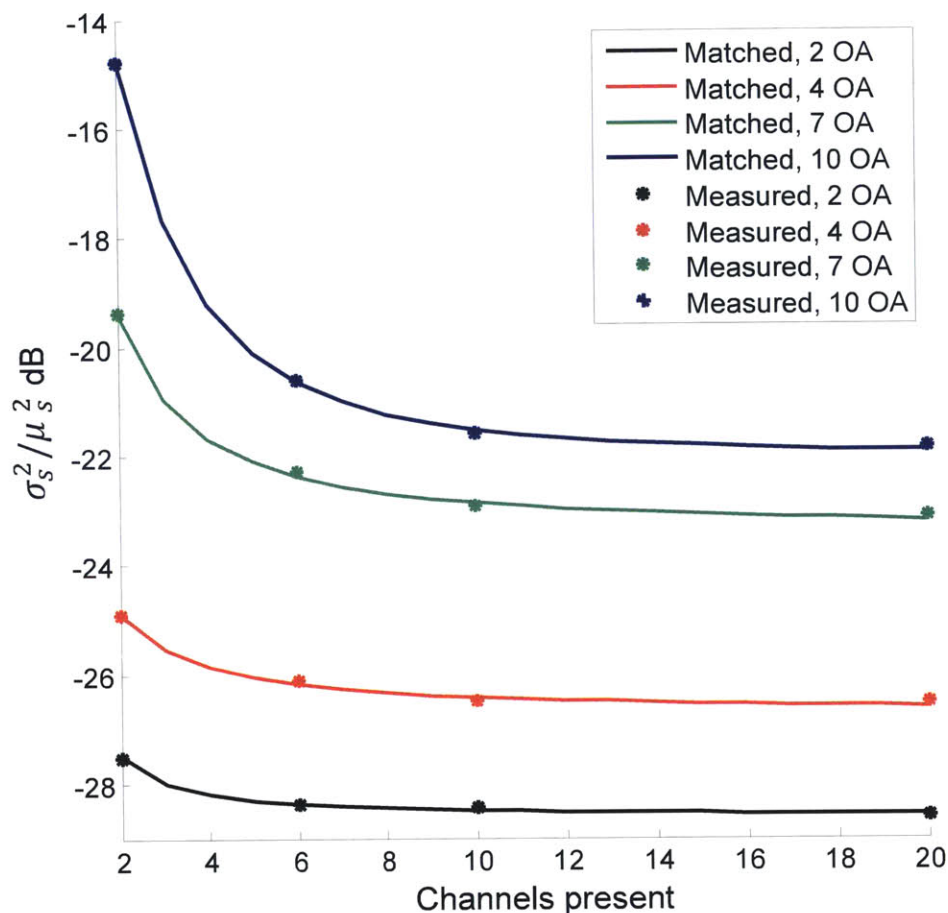
Figure 5-25: Variance/mean-squared of the "1" bit as a function of the channel configuration, matched with a quadratic function of $\frac{1}{N_{ch}-b}$, where $N_{ch}$ is the number of "On" channels in the fiber, and $b$ is a parameter based on the amount of doping of Erbium in the fiber.

Figure 5-26 and 5-27 show the matched signal variance/mean-squared using Model k-G with $lg = P_0/(P_0 + P_{sp})$, which corresponds to the case when the EDFA is under "constant gain control". Letting $P_0 = 5mW$, $B_e = 11GHz$, $B_o = 15GHz$, and $k = [2, 4, 7, 10]$, $\sigma_s^2/\mu_s^2$ can be expressed as a function of fitting parameters $P_{sp}$ and $\sigma_G^2/g^2$. Using the experimental data of $\sigma_s^2/\mu_s^2$, the minimum-mean-square-error estimations of

$P_{sp}$ and $\sigma_G^2/g^2$ are

$$P_{sp} = 3.5 \times 10^{-8} \text{ Watts}, \quad \frac{\sigma_G^2}{g^2} = 0.0006 \qquad \text{for 20 Channels}$$

$$P_{sp} = 3.6 \times 10^{-8} \text{ Watts}, \quad \frac{\sigma_G^2}{g^2} = 0.0007 \qquad \text{for 10 Channels}$$

$$P_{sp} = 3.7 \times 10^{-8} \text{ Watts}, \quad \frac{\sigma_G^2}{g^2} = 0.0008 \qquad \text{for 6 Channels}$$

$$P_{sp} = 4.1 \times 10^{-8} \text{ Watts}, \quad \frac{\sigma_G^2}{g^2} = 0.0025 \qquad \text{for 2 Channels}$$

The model-matched curves agree well with the experimental data for the case with 20 channels, 10 Channels, and 6 Channels, but with larger deviation for the case of 2 channels. This could be explained by the fact that ASE noise span across all channels, and after several amplifiers, ASE noise get accumulated at other "off" channels resembling "on" channels with small signal power. The effect of the amplified-ASE noise at other channels, i.e., the small signals, is more significant when there are only two "On" channels. However, our models only account for ASE noise, i.e., $P_{sp}$, from the channels that are "On", and does not account for the small signal effect from amplified ASE at other "Off" channels. Therefore, the gap between the matched data and the experiment data is more significant for the case of two channels.

However, from our previous analysis of Model k-g in Section 5.5 and Model k-G in Section 5.6, both $lg < 1$ and random gain could result into exponential increase in $\sigma_s^2/\mu_s^2$ with increasing number of amplifiers. The fitting in Fig. 5-26 and 5-27 just offers one possible solution and explanation, and it is not conclusive whether $lg < 1$ or the random gain is the dominating cause. One could investigate the significance of the EDFA random gain by keeping the loss of the fiber span in-between adjacent EDFAs and the average EDFA gain matched, e.g., $lg = 1$, in experiments.
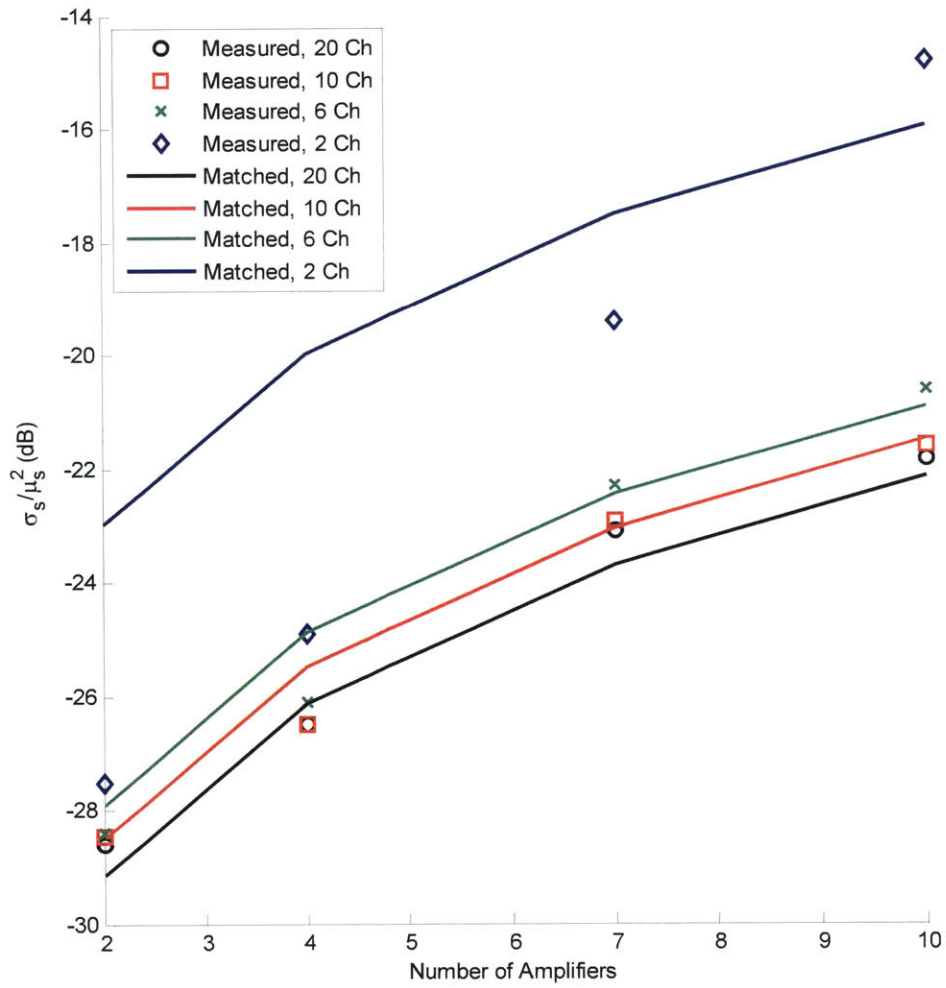
Figure 5-26: Variance/mean-squared of the "1" bit as a function of number of amplifiers, both from experimental data and matched data using Model k-G for the case with constant control, i.e., $lg = P_0/(P_0 + P_{sp})$.
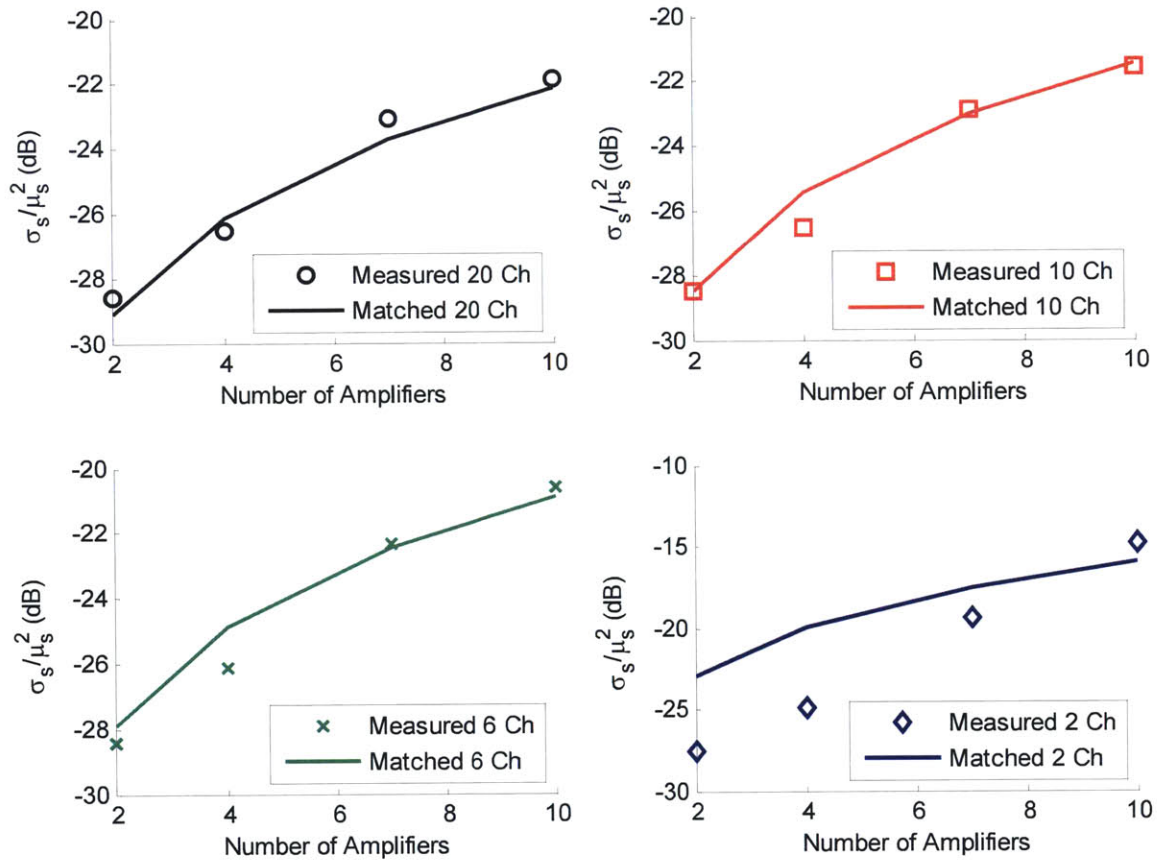
Figure 5-27: Subplots of Variance/mean-squared of the "1" bit as a function of number of amplifiers, both from experimental data and matched data using Model k-G for the case with constant control, i.e., $lg = P_0/(P_0 + P_{sp})$.

# Chapter 6

# Impairment-Aware Scheduling Algorithm and Impairment-Compensation Network Architecture

Both the models and the experiment results in Chap. 5 show that in EDFA-Amplified all-optical networks, channel quality degrades with amplification of more amplifiers, and lightpath merging and diverging at mid-span network from other transactions can cause channel quality variations during the transmission time of one particular transaction. Because both the Meshed and Tunneled Architectures are network logical topologies built on the same physical topology, this EDFA and fast-switching induced channel quality degradation and variations apply to both Architecture M and Architecture T, and add challenges to the scheduling problem of OFS with both architectures.

## 6.1   Impairment-Aware Routing

In Chap. 4, our shortest path algorithm decides the shortest path based on the number of hops of a path to minimize the amount of control traffic to configure switches. With impairment-aware routing, we want to minimize the amount of channel degradation in the shortest path. Assume each optical switch adds $\leq 1dB$ loss which can be ignored comparing tho a total path loss $\geq 20dB$. With this assumption, EDFA is the major source of noise and in the routing algorithm we choose the shortest path to be the path

that has minimum total number of EDFAs. In a well-designed network, the number of EDFAs on a link is highly correlated with the physical link distance, and a longer link usually needs more EDFAs to amplify the signal it carries. Therefore, a path with minimum total number of EDFAs is highly likely to be also the path with the shortest physical distance.

In a network logical topology graph $G(V, E)$, let the weight of each edge $W(i \rightarrow j)$ to be the number of amplifiers on the edge $i \rightarrow j$, then with input $G(V, E)$ and edge weights $W$, the Floyd-Warshall algorithm solves the all-pair shortest paths with running time $\Theta(N_V^3)$, where $N_V$ is the number of nodes in $V$.

## 6.2  Impairment-Aware Scheduled Wavelength Assignment

Chapter 5 shows the number of channels that are "On" in a fiber affects channel quality. Because an end-to-end lightpath across the WAN typically is composed with multiple hops, and in a meshed physical topology, a lightpath may overlap with other lightpaths on one or several hops. Therefore, the turn-on and turn-off of this lightpath may affect the quality of other lightpaths that share a common hop with it. Since in OFS, lightpaths are set up and torn down dynamically, this dynamic status change of lightpaths poses the following challenges on the Scheduling of OFS:

(a) The lightpath quality in the scheduled future can not be determined at the time when the Scheduled Wavelength Assignment is performed, because we do not know whether there will be more transmissions whose requests arrive later but with scheduled transmission times overlapping with the schedule of the current request.

(b) A lightpath that is scheduled into the future affects the channel qualities of the lightpaths that have already been scheduled whose transmission time overlap with the transmission time of the newly scheduled lightpath.

Along a shortest path $P$ with number of hops $H_P$, define $C_{h_i}^P$, a vector of size $\Lambda$ which is the total number of wavelength channels in a fiber, to be the channel configuration of the $i$th hops along path $P$, and $C_{h_i}^P(j) = 1$ if the $j$th wavelength is "On" on the $i$th hops along path $P$. $C_{h_i}^P(j) = 0$, otherwise. Define $f_{\text{BER}} : \{\omega, P, C_{h_i}^P, \text{ for } i = 1, \ldots H_P\} \rightarrow \mathbb{R}$ to be a function that takes into inputs of channels configurations along a shortest path $P$, and outputs the Bit Error Rate of one wavelength $\omega$. Assume such a function can be

found through models of network impairments, then we can verify the channel quality along a scheduled lightpath if the channel configuration is known. Let $th_{\mathrm{BER}}$ to be the maximum BER that can be tolerated. A lightpath is said to be qualified during transmission duration $T$ if its BER is no larger than $th_{\mathrm{BER}}$ for any time in $T$.

### 6.2.1 Impairment-Aware FIFO-EA Algorithm

The Impairment-Aware FIFO-EA Algorithm modifies the FIFO-EA Algorithm to meet the above challenges for the Scheduled Wavelength Assignment Problem along the shortest path. Name this Algorithm the IA-FIFO-EA Algorithm. The IA-FIFO-EA Algorithm still processes the requests in a first-in-first-out manner, and for each request, computes the first available lightpath along the shortest path that satisfies the following conditions:

(a) During the whole scheduled transmission time, the quality of the scheduled lightpath is qualified with input of known future channel configurations.

(b) The set-up and tear-down of this lightpath in the scheduled transmission time will not disqualify any lightpaths that have already been scheduled.

Condition (a) itself does not guarantee the lightpath will be qualified at the scheduled time. This only says that with all that is known about the future at the time of the schedule computation, the lightpath is going to be qualified if there are no future requests that overlap with this schedule. Therefore, together with condition (b) which makes sure scheduling of future requests will not disqualify those that have already been scheduled, we can make sure that a newly scheduled lightpath will be qualified during the scheduled transmission time.

Define $L_{sch}^{P}$ to be the list of transactions that have already been scheduled with at least one hop on $P$ and whose start time is in the future or are currently under transmission. Algorithm 6.1 gives the pseudo-code of the IA-FIFO-EA Algorithm, and Fig. 6-2 illustrates the process of the IA-FIFO-EA Algorithm using a flow chart. Algorithm IA-FIFO-EA repeatedly calls Algorithm FIFO-EA to look for an earliest available wavelength in a future time window starting from $t_0$. The earliest available wavelength $\omega$ is qualified in a scheduled duration $[T_s, T_e]$ if the following two conditions are satisfied.

137

(a) Immediately after the start time and the end time an already-scheduled transaction that falls into $[T_s, T_e]$, the selected $\omega$ is qualified[1].

(b) For any on-going or already scheduled transactions that overlap with $[T_s, T_e]$, the already scheduled wavelengths will not be disqualified by the new transmission assignment.

Condition (a) can be verified by checking the $\omega$ channel quality at the start of the scheduled transmission time and any other times during the transmission when channel configuration changes due to status change of other wavelengths in the same fiber. Figure 6-1 illustrates the possible times when quality of wavelength $\omega$ might change at the change of channel configurations. Condition (b) can be verified by looking into each on-going or already scheduled transaction that overlaps with $[T_s, T_e]$, and checking its quality at times of change of channel configurations during the overlapping time.

**Running Time Analysis of the IA-FIFO-EA Algorithm**

The exact running time depends on the actual traffic arrivals. In the following analysis, we consider the worst case running time in a network that is in steady state. In other words, we use the average occupancies of schedule holders to calculate the effect of the length of the list of transmissions that have already been scheduled, $L_{sh}^P$, and study the worst case where all already-scheduled transmissions overlap with the newly assigned lightpath in both transmission time and transmission path (at least one hop).

Define $R_{\text{IA-FIFO-EA}}^X$ to be the running time of the IA-FIFO-EA Algorithm with Architecture $X$[2], $R_{\text{CHECKQUALITY}}$ to be the running time of the Sub-routine CHECKQUALITY, and $R_{f_{\text{BER}}}$ to be the running time of calculating the BER. Define $N_{\text{FIFO-EA}}^X$ to be the average times that a new schedule needs to be searched by calling FIFO-EA with Architecture X. Because both the start time and end time of a transaction can trigger change of the channel configuration and hence possible channel degradation, in the worst case:

$$R_{\text{IA-FIFO-EA}}^X = N_{\text{FIFO-EA}}^X R_{\text{FIFO-EA}}^X + (|L_{sch}^P| + 1)(R_{\text{CHECKQUALITY}} + \alpha) + R_{f_{\text{BER}}} + \beta \text{ operations}$$

$$(6.1)$$

---

[1]We do not concern fast transients in our algorithm, which are quenched by adiabatic switching and any residual transients are taken care of by the transport layer protocol. Because both starting and ending a transmission cause change of the channel configuration, by verifying channel quality immediately after $T_s$ and $T_e$, we verify channel quality at the new channel configuration.

[2]Architecture X can be Architecture M or Architecture T.
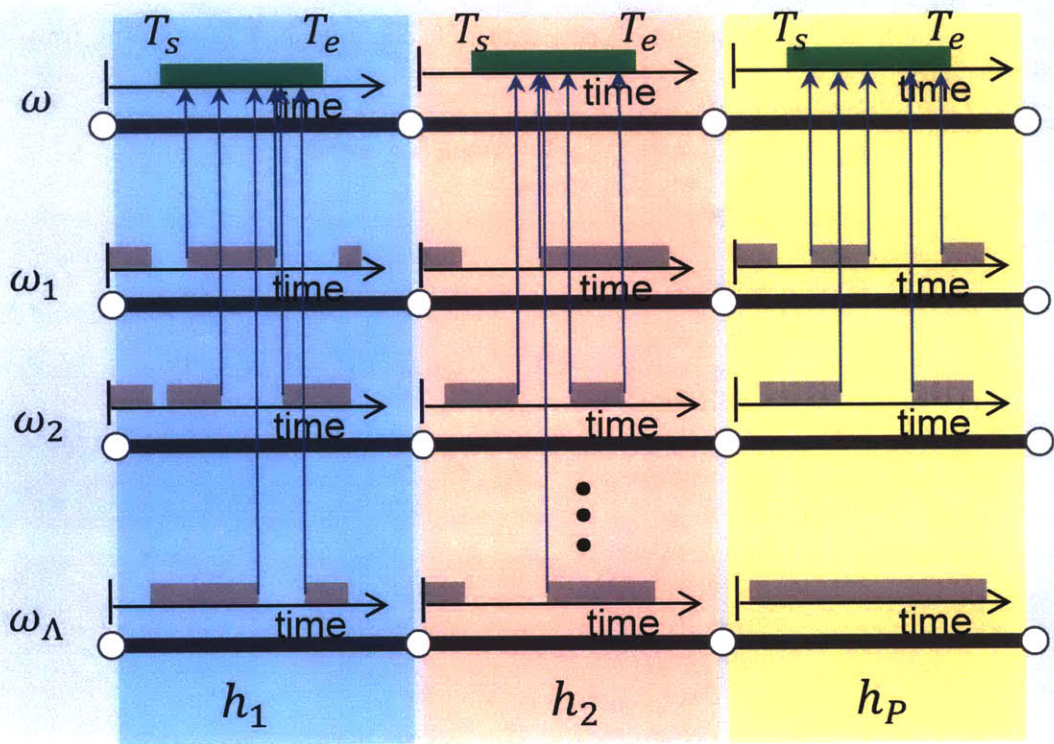
Figure 6-1: Illustration of times of channel configuration change during $[T_s, T_e]$.

**Algorithm 6.1** The IA-FIFO-EA Algorithm

IA-FIFO-EA$(R, P, t_d, L_{sch}^P, t_0)$

1  $t_{\max} \leftarrow$ the latest time among all end times of schedules in $L_{sch}^P$ starting from $t_0$
2  **while** True
3      $\langle \omega, T_s, T_e \rangle \leftarrow$ FIFO-EA$(R, P, t_d, t_0)$
4      **if** $T_s \geq t_{\max}$
5          **if** $f_{\text{BER}}(\omega, P, C_{h_i}^P(T_s)) < \text{th}_{\text{BER}}$
6              **return** $\langle \omega, T_s, T_e \rangle$
7          **else**
8              **return** False
       *//check the quality of $\omega$*
9      $\langle IsQualified, t_{\text{newSearchStart}} \rangle \leftarrow$ CHECKQUALITY$(\omega, P, L_{sch}^P, T_s, T_e)$
10     **if** $IsQualified$ is TRUE
           *//check the quality of all on-going or already-scheduled wavelengths with*
           *//transactions overlapping with $[T_s, T_e]$*
11         **for** each $sch \in L_{sh}^P$
12             $IsPassed \leftarrow$ False
13             **if** $sch$ overlaps with $[T_s, T_e]$ from $T_s'$ to $T_e'$
14                 $\langle IsQualified, t_{\text{newSearchStart}} \rangle \leftarrow$ CHECKQUALITY$(sch.\omega, P, L_{sch}^P, T_s', T_e')$
15                 **if** $IsQualified$ is False
16                     $t_0' \leftarrow t_{\text{newSearchStart}}$
17                     break
18             $IsPassed \leftarrow$ True
19         **if** $IsPassed =$ True
20             **return** $\langle \omega, T_s, T_e \rangle$
21     **else**
22         $t_0' \leftarrow t_{\text{newSearchStart}}$
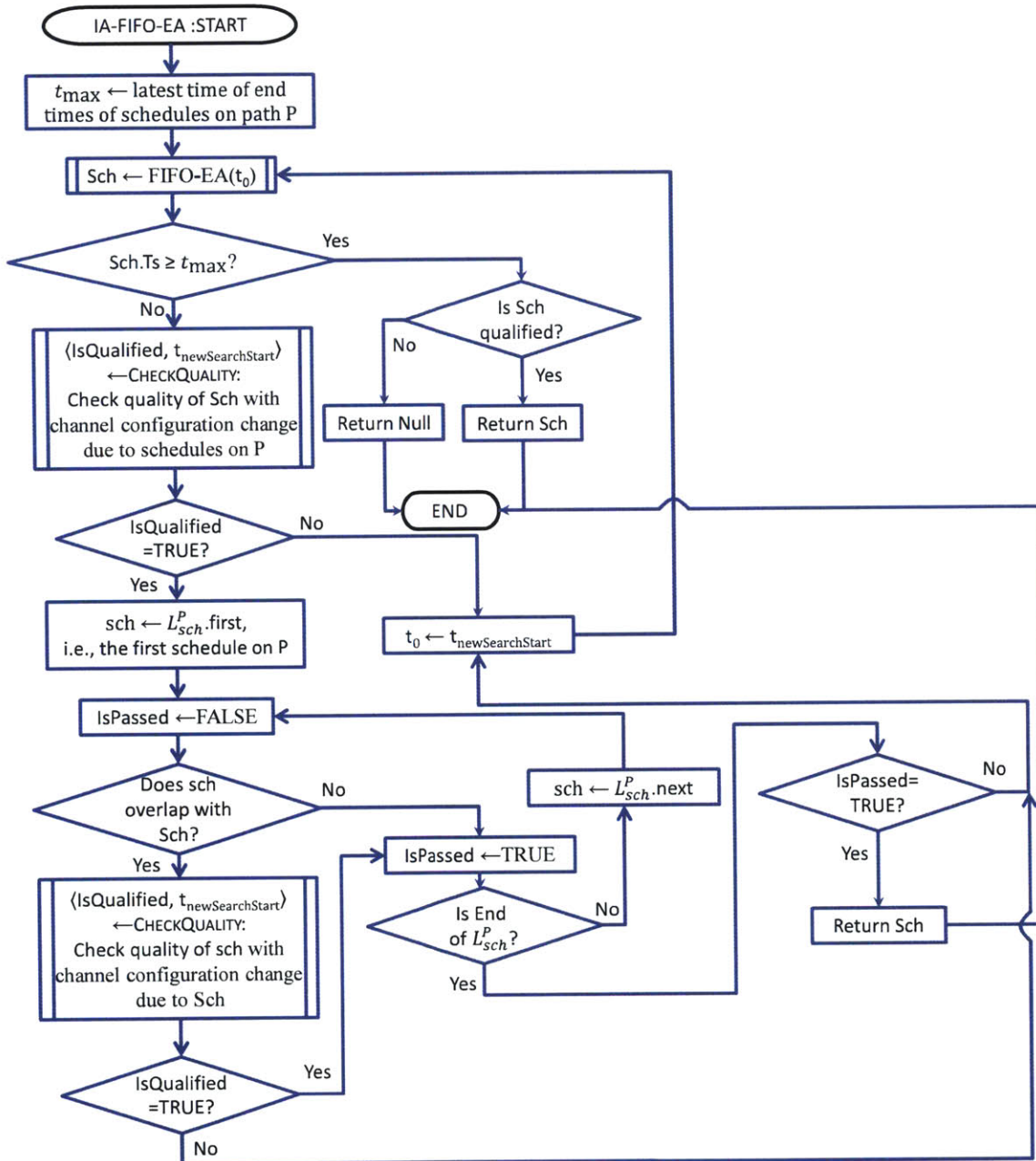
Figure 6-2: Flow chart for the IA-FIFO-EA Algorithm with pseudo-code given by Algorithm 6.1.

**Algorithm 6.2** The CHECKQUALITY Sub-routine

---

CHECKQUALITY$(\omega, P, L_{sch}^P, T_s, T_e)$

1   **if** $f_{\text{BER}}(\omega, P, C_{h_i}^P(T_s)) > \text{th}_{\text{BER}}$

2       $t_0' \leftarrow \min\{t_s, t_e \text{ of } sch \in L_{sh}^P | t_s > T_e, t_e > T_e\}$

3       **return** $\langle \text{False}, t_0' \rangle$

4   **for** each $sch \in L_{sh}^P$

5       **if** $T_s < sch.t_s < T_e$   *//the start time of sch falls into the newly scheduled*
                                             *//transmission time*

6           **if** $f_{\text{BER}}(\omega, P, C_{h_i}^P(sch.t_s)) > \text{th}_{\text{BER}}$   *//wavelength $\omega$ at time $sch.t_s$*
                                                                        *//is not qualified*

7              $t_0' \leftarrow sch.t_s$

8              **return** $\langle \text{False}, t_0' \rangle$

9       **if** $T_s < sch.t_e < T_e$   *//the end time of sch falls into the newly scheduled*
                                             *//transmission time*

10          **if** $f_{\text{BER}}(\omega, P, C_{h_i}^P(sch.t_e)) > \text{th}_{\text{BER}}$   *//wavelength $\omega$ at time $sch.t_e$*
                                                                         *//is not qualified*

11            $t_0' \leftarrow sch.t_e$

12            **return** $\langle \text{False}, t_0' \rangle$

13  **return** $\langle \text{True}, T_s \rangle$

---

where $\alpha$ and $\beta$ capture the running time by comparisons and value assignment operations (in lines 13-17 and 3-6) in the **while** loop of Algorithm 6.1[3], and,

$$R_{\text{CHECKQUALITY}} = (2|L_{sch}^P| + 1)\left(R_{f_{\text{BER}}} + \gamma\right) \quad \text{operations} \tag{6.2}$$

where $\gamma$ captures the running time by comparisons and value assignment operations in lines 5-12 in the **while** loop of Algorithm 6.2[4]. Because on average,

$$|L_{sch}^P| = \bar{H}_P S_X \tag{6.3}$$

where $\bar{H}_P$ is the average number of hops of a path[5], $S_X$ is the average number of

---

[3]The running time of line 1 in Algorithm 6.1 can be ignored comparing to the other operations.

[4]The running time of line 1 in Algorithm 6.2 can be ignored comparing to the other operations.

[5]Note that $\bar{H}_P$ refers to the average number of hops of all source-destination paths in the physical network topology, which is the same as the Meshed logical network topology, and therefore, $\bar{H}_P$ is the same as $\bar{H}_{cp}^M$ and $\bar{H}_{cp}^T$ in Chap. 4.

occupied schedule holders per hop with Architecture X, and,

$$S_M = N_s \rho_{s_M}$$

$$S_T = N_s \rho_{s_T}$$

where $N_s$ is the number of schedule holders per link per fiber. Notice that $\alpha$, $\beta$ and $\gamma$ all include approximately five operations, including comparison, value assignment and returnning values or breaking a loop, we use one parameter $\alpha$ to represent all three of them, and $\alpha \approx 5$. Substituting Eq. (6.3) and (6.2) into Eq. (6.1) leads to,

$$\begin{aligned} R^X_{\text{IA-FIFO-EA}} =& N^X_{\text{FIFO-EA}} R^X_{\text{FIFO-EA}} + \left[ 2(\bar{H}_P S_X)^2 + 3\bar{H}_P S_X + 2 \right] R_{f_{\text{BER}}} \\ &+ \left[ 2(\bar{H}_P S_X)^2 + 4\bar{H}_P S_X + 3 \right] \alpha \quad \text{operations} \end{aligned}$$

From Chap. 4, $R^X_{\text{FIFO-EA}}$ can be approximated as,

$$R^X_{\text{FIFO-EA}} \leq 12 S F \Lambda_0 \bar{H}^X_{dp} \quad \text{operations}$$

where $\bar{H}^X_{dp}$ is the average number of hops of paths over the data plane with $\bar{H}^M_{dp} = 4$ for Architecture M and $\bar{H}^T_{dp} = 1$ for Architecture T. Therefore,

$$\begin{aligned} R^X_{\text{IA-FIFO-EA}} =& N^X_{\text{FIFO-EA}} \cdot 12 S F \Lambda_0 \bar{H}^X_{dp} + \left[ 2(\bar{H}_P S_X)^2 + 3\bar{H}_P S_X + 2 \right] R_{f_{\text{BER}}} \\ &+ \left[ 2(\bar{H}_P S_X)^2 + 4\bar{H}_P S_X + 3 \right] \alpha \quad \text{operations} \end{aligned} \tag{6.4}$$

## 6.2.2 Impairment with Known Worst Case

If we assume the nonlinear effects of fiber on the signal can be ignored, and the fiber noise is dominated by noise generated from the EDFAs, then from the experiment results in Chap. 6, the worst case in terms of channel quality along a given path happens when there is zero wavelength that is "On" in a fiber. With this assumption, the IA-FIFO-EA Algorithm can be simplified and the $f_{\text{BER}}$ can be calculated using models in Chap. 6. If the worst case is when there is no "On" wavelength in a fiber, the major change to the IA-FIFO-EA Algorithm is:

- If a lightpath is qualified based on the condition of the known future of already-scheduled transactions, it is going to stay qualified because during the future transmission along the lightpath, it is only possible to have more "On" wavelengths in the same fiber than the number of wavelengths that were known to be

143

"On" at the time of the schedule computation.

- The already scheduled transactions will not be disqualified by scheduling more transactions that overlap with them in the future.

Therefore, in the new IA-FIFO-EA Algorithm, only the quality of the newly scheduled earliest-available wavelength needs to be verified. Name this Algorithm as the KWC-FIFO-EA Algorithm, with KWC short for "Known Worst Case". The pseudocode of the KWC-FIFO-EA Algorithm is given in Algorithm 6.3, and its process is illustrated using a flow chart in Fig. 6-3.

---

**Algorithm 6.3** The KWC-FIFO-EA Algorithm

---

KWC-FIFO-EA$(R, P, t_d, L^P_{sch}, t_0)$

1   $t_{\max} \leftarrow$ the latest time among all end times of schedules in $L_{sch}$ starting from $t_0$
2   **while** True
3       $\langle \omega, T_s, T_e \rangle \leftarrow$ FIFO-EA$(R, P, t_d, t_0)$
4       **if** $T_s \geq t_{\max}$
5           **if** $f_{\text{BER}}(\omega, P, C^P_{h_i}(T_s)) <$ th$_{\text{BER}}$
6               **return** $\langle \omega, T_s, T_e \rangle$
7           **else**
8               **return** False
        //check the quality of $\omega$
9       $\langle IsQualified, t_{\text{newSearchStart}} \rangle \leftarrow$ CHECKQUALITY$(\omega, P, L^P_{sch}, T_s, T_e)$
10      **if** $IsQualified$ is TRUE
11          **return** $\langle \omega, T_s, T_e \rangle$
12      **else**
13          $t'_0 \leftarrow t_{\text{newSearchStart}}$

---

**Running Time Analysis of the KWC-FIFO-EA Algorithm**

Similarly as Section 6.2.1, we consider the worst case running time in network steady state, using the average occupancies of schedule holders to calculate the effect of the length of the list of transmissions that have already been scheduled, and study the worst case where all already-scheduled transmissions overlap with the newly assigned lightpath in both transmission time and transmission path (at least one hop).

Define $R^X_{\text{KWC-FIFO-EA}}$ to be the running time of the KWC-FIFO-EA Algorithm
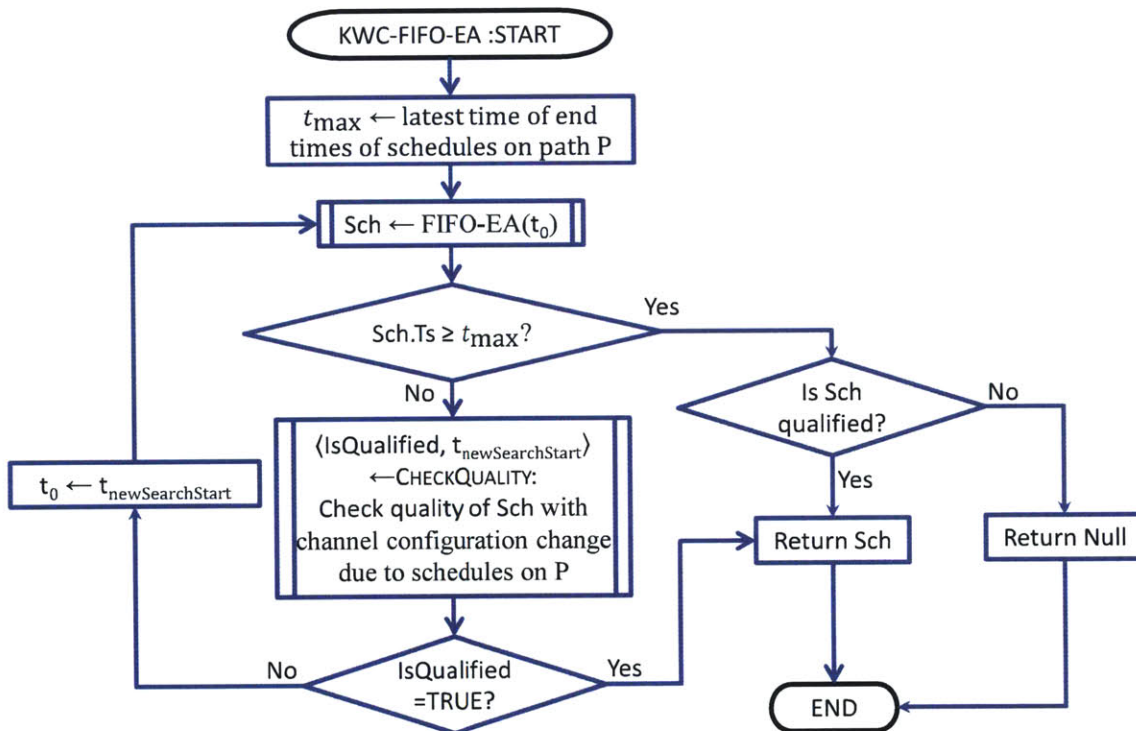
Figure 6-3: Flow chart for the KWC-FIFO-EA Algorithm with pseudo-code given by Algorithm 6.3.

with Architecture $X$. Then, for the worst case,

$$R^X_{\text{KWC-FIFO-EA}} = N^X_{\text{FIFO-EA}} \left( R^X_{\text{FIFO-EA}} + R_{\text{CheckQuality}} + \alpha \right) \quad \text{operations} \quad (6.5)$$

where $\alpha$ captures the running time by comparisons and value assignment operations in the **while** loop of Algorithm 6.3 (lines 4, 5, 10, 11, and 13)[6], and $\alpha \approx 5$. Substituting Eq. (6.3) and (6.2) into Eq. (6.5) leads to,

$$R^X_{\text{KWC-FIFO-EA}} = N^X_{\text{FIFO-EA}} \left[ R^X_{\text{FIFO-EA}} + (2\bar{H}_P S_X + 1) R_{f_{\text{BER}}} + 2(\bar{H}_P S_X + 1)\alpha \right] \quad \text{operations}$$
$$(6.6)$$

With known worst case to be the case when there is zero "On" wavelength in a fiber, we can use models in Chap. 6 to calculate the Bit Error Rate. With EDFA ASE noise measured under different channel configurations and reported through the control plane, and known fiber loss and switch loss, the signal degradation and noise accumulated can be calculated using models in Chap. 6, and the BER can be then estimated using Eq. (5.4) or (5.5). Let $k$ be the total number of amplifiers along a lightpath, using the case when $lg = 1$ and switch loss is negligible for an example, the $f_{\text{BER}}$ can be decomposed into the following steps, each with running time estimated as the number of algebraic operations[7].

1. For bit "One"

   - Calculate the electrical signal power at the end of the lightpath, taking running time $O(1) \approx 3$ operations.

   - Calculate the $sig - sp$ noise in the detection, taking running time $O(1) \approx 10$ operations.

   - Calculate the $sp - sp$ noise in the detection, taking running time $O(1) \approx 11$ operations.

   - Calculate the random gain intensity noise if the gain randomness is known or can be estimated, taking running time $O(\ln k) \approx 60 + 4 \log_2 k$ operations.

2. For bit "Zero"

   - Calculate the electrical signal power at the end of the lightpath, taking running time $O(1) \approx 3$ operations.

---

[6]The running time of line 1 in Algorithm 6.3 can be ignored comparing to the other operations.

[7]Assume shot noise and thermal noise are negligible.

- Calculate the $sp - sp$ noise in the detection, taking running time $O(1) \approx$ 11 operations.

- Calculate the random gain intensity noise if the gain randomness is known or can be estimated, taking running time $O(\ln k) \approx 25 + \log_2 k$ operations.

3. Calculate BER using Eq. (5.5)[8], taking running $O(1) \approx 47$ operations.

Therefore, the running time $R_{f_{\mathrm{BER}}}$ can be estimated to be $R_{f_{\mathrm{BER}}} = 170 + 5 \log_2 k$ operations. Since the total number of amplifiers along a lightpath can be on the orders of 10's, we can approximate $R_{f_{\mathrm{BER}}}$ to be 200 operations. Therefore, Eq. (6.7) can be approximated as,

$$R_{\mathrm{KWC\text{-}FIFO\text{-}EA}}^{X} \approx N_{\mathrm{FIFO\text{-}EA}}^{X} \left( 12 SF \Lambda_0 \bar{H}_{dp}^{X} + 410 \bar{H}_P S_X + 210 \right) \quad \text{operations} \quad (6.7)$$

## 6.3 Processing Power Analysis of Impairment-Aware OFS Scheduling

Because we do not have a model to determine channel quality with the case when fiber nonlinearity can not be ignored and channel quality depends on knowledge of detailed channel configuration, we do not analyze the processing power for the Algorithm IA-FIFO-EA. We focus on the processing power analysis of Algorithm KWC-FIFO-EA, the running time of which has been discussed in Section 6.2.2.

### 6.3.1 Processing Power Analysis of the KWC-FIFO-EA Algorithm

Similar to the analysis in Section 4.6.2, the shortest path routing algorithm is run only when there is a change in the network topology, much less frequently than the KWC- FIFO-EA algorithm. Therefore, we focus on the processing power requirements of the KWC-FIFO-EA Algorithm.

As defined in Chap. 3 and 4, $N_V$ is the number of nodes in the WAN, $\rho$ is the average loading factor of each link, $\Delta$ is the average node degree in the WAN, $N_p$ is the average population size connected to one WAN node, $\lambda$ is the request arrival rate at an end user, and $\bar{T}_d$ is the average transmission time. Then the required processing

---

[8]Assume single-threshold detection is used.

power for Architecture X is $\lambda N_p N_V R^X_{KWC-FIFO-EA}$ for centralized scheduling, and $\lambda N_p R^X_{KWC-FIFO-EA}$ for distributed scheduling. Similar to the analysis in Section 4.6.2, replacing $\lambda N_p$ with $\rho \Delta F \Lambda_0 / \bar{T}_d$, we obtain the total processing power for Architecture X with centralized impairment-aware scheduled wavelength assignment as,

$$P^X_{KWC-cen} = \frac{\rho \Delta F \Lambda_0 N_V R^X_{KWC-FIFO-EA}}{\bar{T}_d}$$

and, with distributed impairment-aware scheduled wavelength assignment,

$$P^X_{KWC-dis} = \frac{\rho \Delta F \Lambda_0 R^X_{KWC-FIFO-EA}}{\bar{T}_d}$$

To obtain numerical values of $P^X_{KWC-cen}$ and $P^X_{KWC-dis}$ for both the average case and the case with peak traffic arrivals, similar as the analysis in Section 4.6.2, we use numerical values in Table 4.4 to calculate the processing power for both the average case $(\rho \bar{D} F \Lambda_0 / \bar{T}_d = 2600)$ and the case of peak traffic arrivals ( $\rho \bar{D} F \Lambda_0 / \bar{T}_d = 9000$). From Fig. (3-10), with these parameters, the blocking probabilities of both Architecture M and T are less than $10^{-3}$, and the schedule holder loading is $\rho_{s_M} = 2.4 \times 10^{-4}$ for Architecture M, and $\rho_{s_T} = 0.14$ for Architecture T. Therefore, the average integer values of $S$ for Architecture M and T are both less than one, hence, $S = 1$, $S_M = 1$ and $S_T = 57$. Table 6.1 summarizes the average processing power for Architecture M and T under both centralized scheduling and distributed scheduling, while Table 6.2 summarizes the peak processing power.

The worst-case processing power is higher for tunneled architecture with Algorithm KWC-FIFO-EA. This is because with Architecture T, because wavelengths are assigned to tunnels quasi-statically, there are more sessions scheduled into the future along a lightpath. Therefore, Algorithm KWC-FIFO-EA needs to check the lightpath quality at more future times when channel configuration changes. However, with distributed scheduling, Algorithm KWC-FIFO-EA can be implemented easily with both Architecture M (1 Intel i7 CPU) and Architecture T (4 Intel i7 CPUs).

Table 6.1: Summary of average processing power for Algorithm FIFO-EA and Algorithm KWC-FIFO-EA.

| Average Processing Power | | FIFO-EA | | KWC-FIFO-EA | |
|---|---|---|---|---|---|
| | | Meshed | Tunneled | Meshed | Tunneled |
| **Centralized Scheduling** | in terms of $N^X_{\text{FIFO-EA}}$ [†] | | | $N^M_{\text{FIFO-EA}} \times 12.27$ GIPS | $N^T_{\text{FIFO-EA}} \times 17.61$ GIPS |
| | numerical Min and Max | 10.7 GIPS | 29.1 MIPS | 12.27 GIPS to 98 GIPS | 17.61 GIPS to 9.58 TIPS |
| | in terms of Intel i7 CPUs[††] | 0.06 CPUs | $1.6 \times 10^{-4}$ CPUs | 0.07 to 0.55 CPUs | 0.1 to 54 CPUs |
| **Distributed Scheduling** | in terms of $N^X_{\text{FIFO-EA}}$ | | | $N^M_{\text{FIFO-EA}} \times 0.2$ GIPS | $N^T_{\text{FIFO-EA}} \times 0.29$ GIPS |
| | numerical Min and Max | 0.18 GIPS | 0.48 MIPS | 0.2 GIPS to 1.6 GIPS | 0.29 GIPS to 157.8 GIPS |
| | in terms of Intel i7 CPUs | 0.001 CPUs | $2.7 \times 10^{-6}$ CPUs | 0.001 to 0.01 CPUs | 0.002 to 0.9 CPUs |

[†] $N^X_{\text{FIFO-EA}}$ ranges from 1 to $2\bar{H}_P S_X$, which is 1 to 8 for Architecture M, and 1 to 544 for Architecture T.

[††] Taking the processing power of the widely available Intel core processor for PC, i7 Extreme Edition 3960X (Hex core) as a reference, it can perform 177.73 GIPS at 3.33 GHz.

Table 6.2: Summary of peak processing power for Algorithm FIFO-EA and Algorithm KWC-FIFO-EA.

| Peak Processing Power | | FIFO-EA | | KWC-FIFO-EA | |
|---|---|---|---|---|---|
| | | Meshed | Tunneled | Meshed | Tunneled |
| **Centralized Scheduling** | in terms of $N^X_{\text{FIFO-EA}}$[†] | | | $N^M_{\text{FIFO-EA}} \times 42.5$ GIPS | $N^T_{\text{FIFO-EA}} \times 60.9$ GIPS |
| | numerical Min and Max | 37.2 GIPS | 0.1 GIPS | 42.5 GIPS to 340 GIPS | 60.9 GIPS to 33.16 TIPS |
| | in terms of Intel i7 CPUs[††] | 0.2 CPUs | $5.6 \times 10^{-4}$ CPUs | 0.24 to 1.9 CPUs | 0.34 to 187.3 CPUs |
| **Distributed Scheduling** | in terms of $N^X_{\text{FIFO-EA}}$ | | | $N^M_{\text{FIFO-EA}} \times 0.7$ GIPS | $N^T_{\text{FIFO-EA}} \times 1$ GIPS |
| | numerical Min and Max | 0.62 GIPS | 1.68 MIPS | 0.7 GIPS to 5.6 GIPS | 1 GIPS to 544 GIPS |
| | in terms of Intel i7 CPUs | 0.004 CPUs | $9.5 \times 10^{-6}$ CPUs | 0.004 to 0.03 CPUs | 0.006 to 3.1 CPUs |

[†] $N^X_{\text{FIFO-EA}}$ ranges from 1 to $2\bar{H}_P S_X$, which is 1 to 8 for Architecture M, and 1 to 544 for Architecture T.

[††] Taking the processing power of the widely available Intel core processor for PC, i7 Extreme Edition 3960X (Hex core) as a reference, it can perform 177.73 GIPS at 3.33 GHz.

# 6.4 Control Traffic Analysis of Impairment-Aware OFS Scheduling

Similar as the control traffic analysis in Section 4.7, the control traffic includes traffic for lightpath requests from users to the Scheduler, control traffic between the Scheduler and the involved switches to set up or tear down lightpaths. In Impairment-Aware OFS Scheduling, the control traffic also includes traffic on reporting the fiber impairments in the network. This EDFA-induced impairment is the same for both Meshed and Tunneled network architecture, but different with respect to whether the scheduling is centralized or distributed. This is because with centralized scheduling, all impairments needs to be reported to one central Scheduler, while with distributed scheduling, all impairments needs to be reported to all Schedulers.

We assume loss, like fiber loss or switch loss, that does not change with channel configurations is already known to the network control unit and, therefore, is not collected nor reported in the control plane. With the assumption that loss caused by fiber nonlinearity is negligible and EDFA-induced noise is the major noise source in the detection, in the control plane, EDFA ASE noise is collected and reported. Assume along one link the ASE noise from different EDFAs is approximately the same, and at the end node of the link, the node is capable of deducing the ASE noise $P_{sp}$ of one EDFA from the accumulated noise of the link[9]. With these assumptions, whenever a transmission starts or ends along a path, nodes along the path need to report the EDFA ASE noise after the change, and,

(a) if detailed channel configurations are needed to calculate BER, for each link, the ASE noise on each channel in the affected fiber needs to be reported together with the new channel configuration;

(b) if worst case is considered, only number of "On" channels and one ASE noise for all channels need to be report for each affected fiber.

For case (a), we use ⟨*link ID, channel configuration, number of amplifiers, $P_{sp}$ for each wavelength*⟩ to capture the impairment. The *link ID* can be an ID related to the fiber which takes 8 bits as discussed in Chap. 4. The *channel configuration* can be $\Lambda$ one-bit identifiers, and the *number of amplifiers* can be represented by an 8-bit Integer. For

---

[9]One way of estimating $P_{sp}$ at the end node of the link is to consider the noise in a wavelength that originated from the start node of that link, whose signal is only corrupted by the EDFAs on that link. Then the $P_{sp}$ can be derived using models in Chap. 5.

$P_{sp}$, we choose to use a 64-bit Long number. Define $L_{IA}$ to be total length of this impairment information with detailed channel configuration, and,

$$L_{IA} = 65\Lambda + 16 \text{ bits}$$

For case (b), the impairment can be captured by $\langle$*link ID, number of amplifiers, number of "On" wavelengths, $P_{sp}\rangle$*. Define $L_{KWC}$ to be total length of the impairment information for known worst case. Using an 8-bit Integer to represent the *number of "On" amplifiers*, then,

$$L_{KWC} = 88 \text{ bits}$$

Again assume TCP/IPv6 is used to transport the impairment control traffic. Notice that the above discussed impairment information is the payload of the IPv6 control packets. Define $L_{IA}^{\text{TCP/IP}}$ or $L_{KWC}^{\text{TCP/IP}}$ to be the total control traffic in one TCP session between one node and the Scheduler for the case of full channel configuration or known worst case, respectively. Then,

$$L_{IA}^{\text{TCP/IP}} = L_{IA}(1 + \alpha_{\text{FEC}}) + 320 \cdot 8 \text{ bits}$$
$$L_{KWC}^{\text{TCP/IP}} = L_{KMC}(1 + \alpha_{\text{FEC}}) + 320 \cdot 8 \text{ bits}$$

where $\alpha_{\text{FEC}}$ denotes the percentage of overhead from the Forward Error Correction relative to the actual length of the payload. Because both starting and ending a transmission trigger impairment reporting of all links along the path $P$ where the transmission happens, using $L_{IA}^{\text{P}}$ and $L_{KWC}^{\text{P}}$ to represent the control traffic triggered by one transmission (both starting and ending of the transmission), then,

$$L_{IA}^{P} = 2L_{IA}^{\text{TCP/IP}} \bar{H}_P \text{ bits}$$
$$L_{KWC}^{P} = 2L_{KWC}^{\text{TCP/IP}} \bar{H}_P \text{ bits}$$

Define $L_{cen}^{Y}$ to be the total control traffic with centralized scheduling, and $L_{dis}^{Y}$ to be the total control traffic with distributed scheduling, where $Y$ can be $IA$ or $KWC$ specifying the two cases of full channel configuration or of the known worst case. Similar as the control traffic analysis in Section 4.7, the total control traffic can be derived from the

traffic demand and the network topology, and,

$$L_{cen}^Y = \frac{\rho \Delta F \Lambda N_V \bar{H}_{cp}}{\bar{T}_d} \cdot L_Y^P$$

$$L_{dis}^Y = \frac{\rho \Delta F \Lambda N_V \bar{H}_{cp}}{\bar{T}_d} \cdot L_Y^P \cdot N_V$$

To obtain numerical values of the control traffic, we use numerical values in Table 4.4 to calculate the control traffic for both the average case ($\rho \bar{D} F \Lambda_0 / \bar{T}_d = 2600$) and the case of peak traffic arrivals ($\rho \bar{D} F \Lambda_0 / \bar{T}_d = 9000$). We obtain numerical values of the additional control traffic for impairment reporting, with the case of average traffic arrivals summarized in Table 6.3 and the case of peak traffic arrivals summarized in Table 6.4.

Table 6.3: Average control traffic for impairment reporting for Algorithm IA-FIFO-EA and KWC-FIFO-EA with both centralized and distributed scheduling.

| Average Control Traffic | IA-FIFO-EA | KWC-FIFO-EA |
|---|---|---|
| Centralized Scheduling | 84.25 Gbps | 13.26 Gbps |
| Distributed Scheduling | 5.06 Tbps | 787.8 Gbps |

Table 6.4: Peak control traffic for impairment reporting for Algorithm IA-FIFO-EA and KWC-FIFO-EA with both centralized and distributed scheduling.

| Peak Control Traffic | IA-FIFO-EA | KWC-FIFO-EA |
|---|---|---|
| Centralized Scheduling | 292.8 Gbps | 45.8 Gbps |
| Distributed Scheduling | 17.5 Tbps | 2.7 Tbps |

Adding the impairment control traffic in Table 6.3 to the control traffic for requests and switch configurations in Table 4.6, we obtain the total control traffic for impairment-aware scheduling with average traffic arrivals summarized in Table 6.5. Similarly, we can obtain the total control traffic for impairment-aware scheduling with peak traffic arrivals summarized in Table 6.6.

Table 6.5: Summary of average control traffic for Algorithm FIFO-EA, IA-FIFO-EA, and KWC-FIFO-EA.

| Average Control Traffic | | FIFO-EA | | IA-FIFO-EA | | KWC-FIFO-EA | |
|---|---|---|---|---|---|---|---|
| | | Meshed | Tunneled | Meshed | Tunneled | Meshed | Tunneled |
| **Centralized Scheduling** | Total traffic | 27.68 Gbps | 12.16 Gbps | 111.93 Gbps | 96.41 Gbps | 40.94 Gbps | 25.4 Gbps |
| | Per-edge | 0.36 Gbps | 0.16 Gbps | 1.45 Gbps | 1.25 Gbps | 0.53 Gbps | 0.33 Gbps |
| | Heaviest edge | 2.66 Gbps | 1.17 Gbps | 10.76 Gbps | 9.27 Gbps | 3.94 Gbps | 2.44 Gbps |
| **Distributed Scheduling** | Total traffic | 312.16 Gbps | 12.16 Gbps | 5.37 Tbps | 5.07 Tbps | 1.1 Tbps | 0.8 Tbps |
| | Per-edge[†] | 4.05 Gbps | 0.16 Gbps | 69.7 Gbps | 65.8 Gbps | 14.39 Gbps | 10.49 Gbps |

[†] With distributed scheduling, since control traffic is sent to all WAN nodes and distributed to all links, the average per-edge load is approximately the heaviest edge load.

Table 6.6: Summary of peak control traffic for Algorithm FIFO-EA, IA-FIFO-EA, and KWC-FIFO-EA.

| Peak Control Traffic | | FIFO-EA | | IA-FIFO-EA | | KWC-FIFO-EA | |
|---|---|---|---|---|---|---|---|
| | | Meshed | Tunneled | Meshed | Tunneled | Meshed | Tunneled |
| **Centralized Scheduling** | Total traffic | 95.9 Gbps | 42.1 Gbps | 388.7 Gbps | 334.9 Gbps | 141.7 Gbps | 87.9 Gbps |
| | Per-edge | 1.26 Gbps | 0.54 Gbps | 5 Gbps | 4.3 Gbps | 1.84 Gbps | 1.15 Gbps |
| | Heaviest edge | 9.2 Gbps | 4.06 Gbps | 37.3 Gbps | 32 Gbps | 13.6 Gbps | 8.5 Gbps |
| **Distributed Scheduling** | Total traffic | 1.1 Tbps | 42.1 Gbps | 18.6 Tbps | 17.6 Tbps | 3.8 Tbps | 2.76 Tbps |
| | Per-edge[†] | 14 Gbps | 0.54 Gbps | 251 Gbps | 230 Gbps | 49.8 Gbps | 36.4 Gbps |

† With distributed scheduling, since control traffic is sent to all WAN nodes and distributed to all links, the average per-edge load is approximately the heaviest edge load.

Detailed channel configuration information leads to huge control traffic for link impairment reporting. With distributed scheduling, for the case of peak traffic arrivals, the per-edge control traffic is 251 Gbps for Architecture M, and 230 Gbps for Architecture T, which is huge control plane burden. However, if the network is designed to stay away from the fiber nonlinearity region[10], with Algorithm KWC-FIFO-EA and distributed scheduling, the per-edge control traffic is reduced to manageable amount, 49.8 Gbps for Architecture M (with 72% for link impairment reporting and 28% for lightpath setup/tear-down), and 36.4 Gbps for Architecture T (with 99% for link impairment reporting and 1% for lightpath setup/tear-down).

## 6.5    Impairment-Compensation Network Architecture

In this section, we propose a new network architecture, where the network is designed such that the worst case of channel quality is when there is no "On" wavelength in a fiber[11]. To ensure during the worst case when there is no other "On" wavelength the lightpath is still qualified or to increase the reach of a lightpath, a dummy laser is used to inject optical power to each fiber going out of each switch (Fig. 6-4). The power of the laser can be controlled in a way as shown in Fig. 6-5 so that the optical power in any fiber is always above a certain minimum level which ensures the worst case or the case with a longer reach is qualified. The total number of lasers equal to the number of fibers in the network, which is, $2N_E F$, where $N_E$ is the number of bidirectional edges in the network and $F$ is the number of fibers for each edge. Note that even with this Impairment-Compensation network architecture, link impairment should still be reported.

---

[10]This could be achieved by lowering the input optical power, and allowing a higher BER using Forward Error Correction.

[11]The design can limit the effect of fiber nonlinearities induced cross-channel interference so that EDFA-induced noise dominates.

156
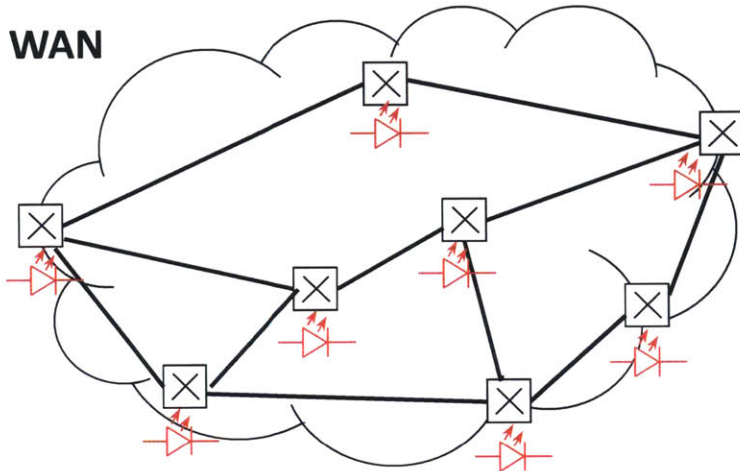
Figure 6-4: New WAN architecture with a dummy laser for each fiber out of a switch.



Dummy laser power
$$\min\bigl(0, P_{\min} - P_{\text{switch}}\bigr)$$

Switch power

Dummy laser power

Total power into fiber
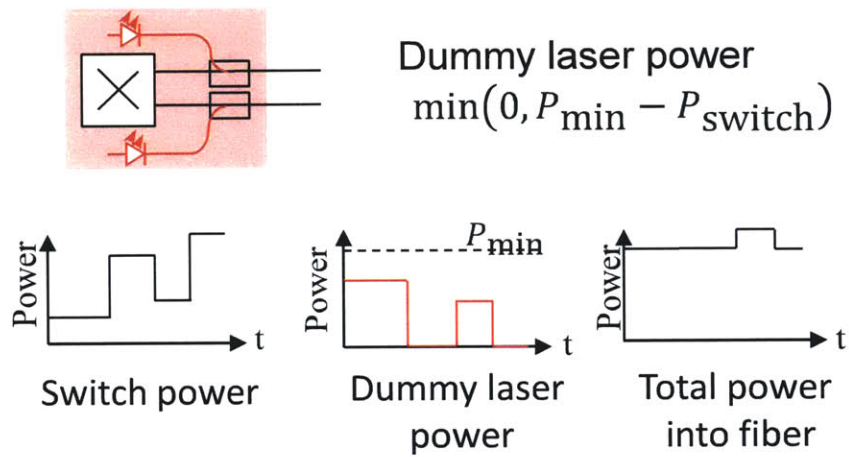
Figure 6-5: Illustration of using dummy lasers to make sure the total optical power in a fiber is maintained above a minimum value.

# Chapter 7

# Conclusion

In this thesis, we have addressed the network management and control aspects of flow-switched optical networks, and designed a network architecture that leads to both a scalable control plane and an efficient data plane.

Systematically, we have given an overview of the essential network management and control entities and the corresponding functionalities. We focused on the scheduling problem of OFS because its processing power and generated control traffic increase with traffic demand, network size, and are closely correlated with data network architecture, while other routine maintenance type network management and control functionalities contribute either a fixed amount or negligibly to the total efforts. We consider two possible Wide Area Network architectures: meshed or tunneled, and developed a unified model for data plane performance to provide a common platform for the performance comparison of the control plane. We have also developed a framework to analyze network management and control efforts from the perspectives of processing complexity and control traffic as functions of network architecture, traffic demand, and network resources. The presence of switching-induced physical-layer impairments in EDFA-amplified optical links creates more challenges to network management and control. To incorporate the effects of the physical-layer impairments into the scheduling algorithm to guarantee the quality of the scheduled lightpath, we have developed models for quality of EDFA-amplified optical links. We have developed impairment-aware scheduling algorithms for two cases, a) the worst case of channel quality happens when there is no "On" channel in a fiber, and b) full channel configurations of a fiber are needed to determine the quality of one channel and the worst case is not known. Finally, we have proposed an impairment-compensation network architecture which

employs dummy lasers at each switch to guarantee the worst-case channel quality or to increase the reach of a lightpath.

Our analysis of the data plane performance has shown that with aggregation of at least two wavelengths of traffic and allowing about two transactions per wavelength to be scheduled into the future, the tunneled architecture provides comparable data plane performance as the meshed architecture. However, with no physical-layer impairments, the tunneled architecture reduces the processing power and the control traffic by orders of magnitude. Taking the case of peak traffic arrivals with distributed scheduling for an example, the peak processing complexity is reduced from $\sim$0.6 GIPS to $\sim$1.7 MIPS, and the peak control traffic per edge is reduced from $\sim$14 Gbps to $\sim$0.5 Gbps. With impairment-aware scheduling, our results show that if detailed channel configuration information is reported in the control plane, link impairment updates lead to huge control traffic, e.g., $\sim$250 Gbps per-edge peak control traffic with distributed scheduling over the meshed architecture. However, if the worst case of impairment is known to be when there is no "On" channel in a fiber, using the tunneled architecture, both the processing power and the control traffic are reduced to manageable amounts. For example, the scheduling can be handled by 1 to 4 Intel i7 CPUs, and the per-edge peak control traffic is $\sim$36 Gbps.

Below are possible directions for future work.

1. In our modeling of the data plane performance, we assumed a symmetric network topology graph. When the actual network topology graph is not symmetric, one could study to which extent our model still provides a good approximation of the data plane performance.

2. For scheduling over the meshed architecture, due to mid-span merging and diverging traffic, gaps between adjacent schedules over a channel could be unavoidable, leading to possible inefficient network resource utilization. In addition, the expected number of occupied schedule holders over a path do not necessarily indicate the expected delay of a new request along that path with the meshed architecture. New models are needed to investigate the amount of gaps and the delay for scheduling over the meshed architecture.

3. One could investigate the significance of EDFA random gain noise by keeping the loss of the fiber span between adjacent EDFAs and the average EDFA gain matched in experiments.

# Appendix A

# Derivations of Equations and Theorems in Chapter 3

## A.1  Proof of Theorem 1

*Proof.* By induction. For $i = 2$,

$$
\begin{aligned}
\rho_{s2} &= (1 - P_{sl} + P_{sn}P_{sl})\rho_{s1} + P_{sn}(1 - \rho_{s1}) \\
&= \frac{(1 - P_{sl} + P_{sn}P_{sl})P_{sn}}{P_{sn} + P_{sl} - P_{sn}P_{sl}} + \frac{P_{sn}(P_{sl} - P_{sn}P_{sl})}{P_{sn} + P_{sl} - P_{sn}P_{sl}} \\
&= \frac{P_{sn}}{P_{sn} + P_{sl} - P_{sn}P_{sl}} \\
&= \rho_s.
\end{aligned}
$$

Assume $\rho_{sk} = \rho_s$, then for $i = k + 1$,

$$
\begin{aligned}
\rho_{s(k+1)} &= (1 - P_{sl} + P_{sn}P_{sl})\rho_{sk} + P_{sn}(1 - \rho_{sk}) \\
&= \frac{(1 - P_{sl} + P_{sn}P_{sl})P_{sn}}{P_{sn} + P_{sl} - P_{sn}P_{sl}} + \frac{P_{sn}(P_{sl} - P_{sn}P_{sl})}{P_{sn} + P_{sl} - P_{sn}P_{sl}} \\
&= \frac{P_{sn}}{P_{sn} + P_{sl} - P_{sn}P_{sl}} \\
&= \rho_s.
\end{aligned}
$$

$\square$

# Appendix B

# Detection of Two Gaussian Random Variables with Non-equal Means and Variations

In this appendix, we derive the optimal threshold detector for the hypothesis testing on whether the received $X$ is "1" or "0" using techniques from [37].

## B.1   Optimal Two-threshold Detector

Assume under hypothesis 1, $X$ is a Gaussian random variable with mean $\mu_s$, and variance $\sigma_s^2$, that is,

$$H_1: \quad X \sim \text{Gaussian}(\mu_s, \sigma_s^2)$$

Assume under hypothesis 0, $X$ is a Gaussian random variable with mean $\mu_n$, and variance $\sigma_n^2$, that is,

$$H_0: \quad X \sim \text{Gaussian}(\mu_n, \sigma_n^2)$$

In addition, we assume,

$$\mu_n < \mu_s$$
$$\sigma_n^2 < \sigma_s^2$$

Therefore, the distribution of $X$ is,

$$H_1: \quad f_{x|H_1}(x) = \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left[-\frac{(x-\mu_s)^2}{2\sigma_s^2}\right] \tag{B.1}$$

$$H_0: \quad f_{x|H_0}(x) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left[-\frac{(x-\mu_n)^2}{2\sigma_n^2}\right] \tag{B.2}$$

Solving $f_{x|H_1}(x) = f_{x|H_0}(x)$ leads to two crossing points of the two Gaussian distributions, see Fig. B-1, and,

$$x_1^* = a + \sqrt{a^2 - b} \tag{B.3}$$

$$x_2^* = a - \sqrt{a^2 - b} \tag{B.4}$$

$$a = \frac{\mu_n \sigma_s^2 - \mu_s \sigma_n^2}{\sigma_s^2 - \sigma_n^2}$$

$$b = \frac{\sigma_s^2 \mu_n^2 - \sigma_n^2 \mu_s^2 + 2\sigma_s^2 \sigma_n^2 \log\left(\frac{\sigma_n}{\sigma_s}\right)}{\sigma_s^2 - \sigma_n^2}$$

The optimal threshold detector is a two-threshold detector with thresholds $x_1^*$ and $x_2^*$, and,

$$H_1 \text{ is true, if} \quad X < x_2^*, \text{ or } X > x_1^*$$

$$H_0 \text{ is true, if} \quad x_2^* < X < x_1^*$$

To see why the above is true, consider the error probability using a two-threshold detector with thresholds at $x_1$ and $x_2$, and $x_2 < \mu_n < x_1$, and,

$$P_e = P(H_0)Pr\{X < x_2, \text{ or, } X > x_1 | H_0\} + P(H_1)Pr\{x_2 < X < x_1 | H_1\} \tag{B.5}$$

Since "1" and "0" bits are equally likely, $P(H_1) = P(H_0) = 0.5$. Define $\phi_i(x)$ as the cumulative probability distribution of the Gaussian random variable under hypothesis $H_i$. Then Eq. (B.5) can be re-written as,

$$P_e = \frac{1}{2}\left(\phi_0(x_2) + 1 - \phi_0(x_1)\right) + \frac{1}{2}\left(\phi_1(x_1) - \phi_1(x_2)\right)$$

$$= \frac{1}{2}\underbrace{\left(1 - \phi_0(x_1) + \phi_1(x_1)\right)}_{s(x_1)} + \frac{1}{2}\underbrace{\left(\phi_0(x_2) - \phi_1(x_2)\right)}_{g(x_2)}$$

164
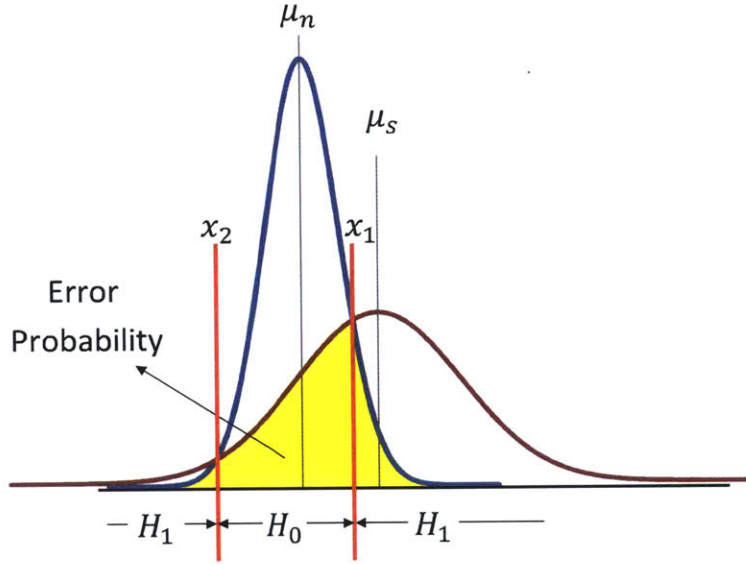
Figure B-1: Two-threshold detection model of two Gaussian random variables with non-equal means and variances.

Note that as long as $x_2 < \mu_n < x_1$ is maintained, $P_e$ can be optimized by choosing the $x_1^*$ and $x_2^*$ independently. Therefore,

$$\min_{x_2 < \mu_n < x_1} P_e = \min_{x_1 > \mu_n} s(x_1) + \min_{x_2 < \mu_n} g(x_2)$$

The first derivatives of $s(x_1)$ and $g(x_2)$ w.r.t. $x_1$ and $x_2$, respectively, are

$$s'(x_1) = f_{x|H_1}(x_1) - f_{x|H_0}(x_1)$$
$$g'(x_2) = f_{x|H_0}(x_2) - f_{x|H_1}(x_2)$$

Setting the above first derivatives equal to zero, leads to

$$f_{x|H_1}(x_1) = f_{x|H_0}(x_1)$$
$$f_{x|H_0}(x_2) = f_{x|H_1}(x_2)$$

Solving the above equations, we have obtained $x_1^*$ in Eq. (B.3) and $x_2^*$ in Eq. (B.4), and $x_2^* < \mu_n < x_1^*$.

165

The second derivatives of $s(x_1)$ and $g(x_2)$ w.r.t. $x_1$ and $x_2$, respectively, are

$$s''(x_1) = -\frac{(x_1 - \mu_s)^3}{\sqrt{2\pi}\sigma_s^3} + \frac{(x_1 - \mu_n)^3}{\sqrt{2\pi}\sigma_n^3}$$

$$g''(x_2) = \frac{(x_2 - \mu_s)^3}{\sqrt{2\pi}\sigma_s^3} - \frac{(x_2 - \mu_n)^3}{\sqrt{2\pi}\sigma_n^3}$$

At $x_1 = x_1^*$,

$$s''(x_1^*) = -\frac{(x_1^* - \mu_s)^3}{\sqrt{2\pi}\sigma_s^3} + \frac{(x_1^* - \mu_n)^3}{\sqrt{2\pi}\sigma_n^3}$$

$$= \frac{\sigma_s^3(x_1^* - \mu_n)^3 - \sigma_n^3(x_1^* - \mu_s)^3}{\sqrt{2\pi}\sigma_n^3\sigma_s^3} \tag{B.6}$$

$$> \frac{\sigma_s^3\left[(x_1^* - \mu_n)^3 - (x_1^* - \mu_s)^3\right]}{\sqrt{2\pi}\sigma_n^3\sigma_s^3} \tag{B.7}$$

$$> 0 \tag{B.8}$$

where Eq. (B.6) $\Rightarrow$ Eq. (B.7) because $\sigma_s^2 > \sigma_n^2$, and Eq. (B.7) $\Rightarrow$ Eq. (B.8) because $x_1^* > \mu_n$ and $\mu_n < \mu_s$. Similarly, at $x_2 = x_2^*$,

$$g''(x_2^*) > 0$$

Therefore,

$$\min_{x_1 > \mu_n} s(x_1) = s(x_1^*)$$

$$\min_{x_2 < \mu_n} g(x_2) = g(x_2^*)$$

and,

$$\min_{x_2 < \mu_n < x_1} P_e = \frac{1}{2}\left(1 - \phi_0(x_1^*) + \phi_1(x_1^*)\right) + \frac{1}{2}\left(\phi_0(x_2^*) - \phi_1(x_2^*)\right) \tag{B.9}$$

In communication systems, error probability is often expressed as functions of the Q-function. The Q-function is the tail probability of the standard normal distribution, that is,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{x^2}{2})\mathrm{d}x = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right) \tag{B.10}$$

Expressing Eq. (B.9) using Q-function, we obtain the minimum error probability

of the optimal two-threshold detector, and,

$$(P_e)_{\min} = \frac{1}{2}\left[Q\left(\frac{\mu_s - x_1^*}{\sigma_s}\right) + Q\left(\frac{x_1^* - \mu_n}{\sigma_n}\right)\right] - \frac{1}{2}\left[Q\left(\frac{\mu_s - x_2^*}{\sigma_s}\right) - Q\left(\frac{\mu_n - x_2^*}{\sigma_n}\right)\right]$$

$$(B.11)$$

## B.2 Single-threshold Detector Approximation

Consider a single threshold detector with threshold at $x_1^*$ (see Fig. B-2), and

$$H_1 \text{ is True, if } X > x_1^*$$

$$H_0 \text{ is True, if } X < x_1^*$$
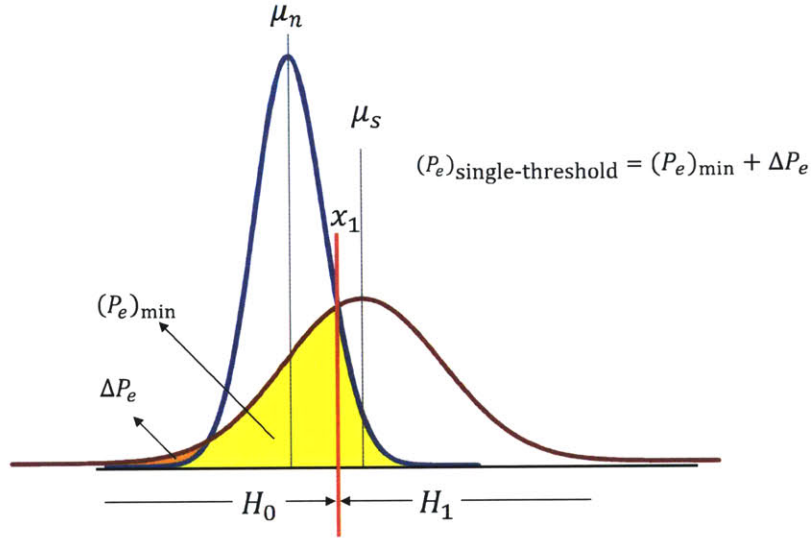


Figure B-2: Single-threshold detection model of two Gaussian random variables with non-equal means and variances.

Therefore, the error probability is:

$$(P_e)_{\text{single-threshold}} = P(H_1)Pr\left(\frac{X - \mu_s}{\sigma_s} < \frac{x_1^* - \mu_s}{\sigma_s}\Big|H_1\right) + P(H_0)Pr\left(\frac{X - \mu_n}{\sigma_n} > \frac{x_1^* - \mu_n}{\sigma_n}\Big|H_0\right)$$

$$= \frac{1}{2}\left[Q\left(\frac{\mu_s - x_1^*}{\sigma_s}\right) + Q\left(\frac{x_1^* - \mu_n}{\sigma_n}\right)\right]$$

The difference of the error probability between the optimal two-threshold detector

and the single-threshold detector is

$$\Delta P_e \triangleq (P_e)_{\text{single-threshold}} - (P_e)_{\min}$$

$$= \frac{1}{2}\left[ Q\left(\frac{\mu_s - x_2^*}{\sigma_s}\right) - Q\left(\frac{\mu_n - x_2^*}{\sigma_n}\right)\right]$$

From [12],

$$Q(x) \approx \frac{1}{12}\exp(-\frac{x^2}{2}) + \frac{1}{4}\exp(-\frac{2}{3}x^2) \quad x > 0$$

Therefore, $Q(x)$ decreases exponentially with increase of $x$, and,

$$\frac{\Delta P_e}{(P_e)_{\text{single-threshold}}} = \frac{Q\left(\frac{\mu_s - x_2^*}{\sigma_s}\right) - Q\left(\frac{\mu_n - x_2^*}{\sigma_n}\right)}{Q\left(\frac{\mu_s - x_1^*}{\sigma_s}\right) + Q\left(\frac{x_1^* - \mu_n}{\sigma_n}\right)}$$

$$< \frac{Q\left(\frac{\mu_s - x_2^*}{\sigma_s}\right)}{Q\left(\frac{\mu_s - x_1^*}{\sigma_s}\right)}$$

$$\approx \frac{\frac{1}{12}\exp\left(-\frac{1}{2}\left(\frac{\mu_s - x_2^*}{\sigma_s}\right)^2\right) + \frac{1}{4}\exp\left(-\frac{2}{3}\left(\frac{\mu_s - x_2^*}{\sigma_s}\right)^2\right)}{\frac{1}{12}\exp\left(-\frac{1}{2}\left(\frac{\mu_s - x_1^*}{\sigma_s}\right)^2\right) + \frac{1}{4}\exp\left(-\frac{2}{3}\left(\frac{\mu_s - x_1^*}{\sigma_s}\right)^2\right)}$$

Since

$$\frac{\Delta P_e}{(P_e)_{\min}} = \frac{1}{\frac{1}{\frac{\Delta P_e}{(P_e)_{\text{single-threshold}}}} - 1}$$

$\Delta P_e$ is negligible relative to $(P_e)_{\min}$ if $x_2^* \ll x_1^2$, that is,

$$(\mu_n \sigma_s^2 - \mu_s \sigma_n^2)^2 \gg \sigma_s^2 \mu_n^2 - \sigma_n^2 \mu_s^2 + 2\sigma_s^2 \sigma_n^2 \log\left(\frac{\sigma_n}{\sigma_s}\right)$$

# Appendix C

# Derivations of Equations in Chapter 5

## C.1  Derivation of $\sigma^2_{IG_{sig}}$ in Eq. (5.24)

*Proof.* The gain intensity noise from the signal is,

$$\sigma^2_{IG_{sig}} = \mathrm{Var}\left(\left|E_0(lg)^{\frac{k}{2}}\left(\prod_{i=1}^{k}\frac{G_i}{g}\right)^{1/2}\right|\right)$$

$$= P_0(lg)^k\left(E\left[\prod_{i=1}^{k}\frac{G_i}{g}\right] - E\left[\prod_{i=1}^{k}\left(\frac{G_i}{g}\right)^{1/2}\right]^2\right)$$

$$= P_0(lg)^k\left[1 - \left(\frac{E[\sqrt{G}]^2}{g}\right)^k\right]$$

$$= P_0(lg)^k\left[1 - \left(\frac{g - \mathrm{Var}(\sqrt{G})}{g}\right)^k\right]$$

$$= P_0(lg)^k\left[1 - \left(1 - \frac{\mathrm{Var}(\sqrt{G})}{g}\right)^k\right] \qquad \square$$

Note that in the above calculation, for simplicity reasons, real field values are considered.

## C.2 Derivation of $\sigma^2_{IG_{sp}}$ in Eq. (5.25)

*Proof.* The gain intensity noise from the accumulated ASE is,

$$\sigma^2_{IG_{sp}} = \text{Var}\left(\left\|\sum_{i=1}^{k-1} E_{sp}\left(\prod_{j=i+1}^{k}\frac{G_j lg}{g}\right)^{\frac{1}{2}} + E_{sp}\right\|\right)$$

$$= \text{Var}\left(\left\|\sum_{i=1}^{k-1} E_{sp}\left(\prod_{j=i+1}^{k}\frac{G_j lg}{g}\right)^{\frac{1}{2}}\right\|\right)$$

$$= P_{sp}\left\{E\left[\left[\sum_{i=1}^{k-1}\prod_{j=i+1}^{k}\left(\frac{G_j lg}{g}\right)^{\frac{1}{2}}\right]^2\right] - \left[E\left[\sum_{i=1}^{k-1}\prod_{j=i+1}^{k}\left(\frac{G_j lg}{g}\right)^{\frac{1}{2}}\right]\right]^2\right\}$$

$$= P_{sp}\left\{E\left[\sum_{i=1}^{k-1}\prod_{j=i+1}^{k}\frac{G_j lg}{g} + 2\sum_{i=1}^{k-2}\sum_{m=i+1}^{k-1}\prod_{j=i+1}^{k}\left(\frac{G_j lg}{g}\right)^{\frac{1}{2}}\prod_{n=m+1}^{k}\left(\frac{G_m lg}{g}\right)^{\frac{1}{2}}\right] - \left[\sum_{i=1}^{k-1}E\left[\prod_{j=i+1}^{k}\left(\frac{G_j lg}{g}\right)^{\frac{1}{2}}\right]\right]^2\right\}$$

$$= P_{sp}\left\{\sum_{i=1}^{k-1}E\left[\prod_{j=i+1}^{k}\frac{G_j lg}{g}\right] - \sum_{i=1}^{k-1}E\left[\prod_{j=i+1}^{k}\left(\frac{G_j lg}{g}\right)^{\frac{1}{2}}\right]^2\right\}$$

$$= P_{sp}\sum_{i=1}^{k-1}\left(\prod_{j=i+1}^{k}E\left[\frac{G_j lg}{g}\right] - \prod_{j=i+1}^{k}E\left[\sqrt{\frac{G_j lg}{g}}\right]^2\right)$$

$$= P_{sp}\sum_{i=1}^{k-1}\left((lg)^{k-i} - \prod_{j=i+1}^{k}lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\right)$$

$$= P_{sp}\sum_{i=1}^{k-1}\left((lg)^{k-i} - \left[lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\right]^{k-i}\right)$$

$$= P_{sp}\left\{\frac{lg\left(1-(lg)^{k-1}\right)}{1-lg} - \frac{lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\left[1 - \left(lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\right)^{k-1}\right]}{lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)}\right\}$$

$$= \begin{cases} P_{sp}\left\{(k-1) - \frac{g}{\text{Var}(\sqrt{G})}\left[1 - \left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)^{k-1}\right]\right\} & \text{if } lg = 1 \\[3mm] P_{sp}\left\{\frac{lg\left(1-(lg)^{k-1}\right)}{1-lg} - \frac{lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\left[1 - \left(lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)\right)^{k-1}\right]}{1-lg\left(1 - \frac{\text{Var}(\sqrt{G})}{g}\right)}\right\} & \text{if } lg \neq 1, lg \neq \frac{1}{1 - \frac{\text{Var}(\sqrt{G})}{g}} \quad \square \\[3mm] P_{sp}\left\{\frac{lg\left[(lg)^{k-1}-1\right]}{lg-1} - (k-1)\right\} & \text{if } lg = \frac{1}{1 - \frac{\text{Var}(\sqrt{G})}{g}} \end{cases}$$

Note that in the above calculation, for simplicity reasons, real field values are considered.

## C.3 Derivation of $\sigma^2_{IG_{sig-sig}}$ in Eq. (5.29)

*Proof.*

$$\sigma^2_{IG_{sig-sig}} = R^2 P_0^2 (lg)^{2k} \mathrm{Var}\left(\prod_{i=1}^{k} \frac{G_i \beta_i}{g}\right)$$

$$= R^2 P_0^2 (lg)^{2k} \left(E\left[\prod_{i=1}^{k}\left(\frac{G_i}{g}\right)^2\right] - E\left[\prod_{i=1}^{k}\frac{G_i}{g}\right]^2\right)$$

$$= R^2 P_0^2 (lg)^{2k} \left(E\left[\left(\frac{G_i}{g}\right)^2\right]^k - 1\right)$$

$$= R^2 P_0^2 (lg)^{2k} \left(\left(\frac{g^2 + \sigma_G^2}{g^2}\right)^k - 1\right)$$

$$= R^2 P_0^2 (lg)^{2k} \left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^k - 1\right] \qquad \square$$

## C.4 Derivation of $\sigma^2_{IG_{sig-sp}}$ in Eq. (5.30)

*Proof.* Define $(\sigma_I^2)_{sig-sp,i}$ to be the intensity noise from signal beating with the $i$th noise term. The corresponding photocurrent is

$$I_{sig-sp,i} = 2R \int \left[P_0 \cdot lG_1 \cdot lG_2 \cdots lG_i \cdot S_{sp}\left(lG_{i+1} \cdot lG_{i+2} \cdots lG_k\right)^2\right]^{\frac{1}{2}} d\omega_n$$

Therefore,

$$(\sigma_I^2)_{sig-sp,i} = \mathrm{Var}\left(I_{sig-sp,i}\right)$$

$$= 4R^2 P_0 S_{sp} B_e \cdot \mathrm{Var}\left(\prod_{j=1}^{i}\left(\frac{G_j lg}{g}\right)^{\frac{1}{2}} \prod_{m=i+1}^{k} \frac{G_m lg}{g}\right)$$

$$= 4R^2 P_0 S_{sp} B_e (lg)^{2k-i} \cdot \left\{E\left[\left(\prod_{j=1}^{i}\left(\frac{G_i}{g}\right)^{\frac{1}{2}} \prod_{m=i+1}^{k} \frac{G_m}{g}\right)^2\right] - E\left[\prod_{j=1}^{i}\left(\frac{G_i}{g}\right)^{\frac{1}{2}} \prod_{m=i+1}^{k} \frac{G_m}{g}\right]^2\right\}$$

$$= 4R^2 P_0 S_{sp} B_e (lg)^{2k-i} \cdot \left\{E\left[\prod_{m=i+1}^{k} \frac{G_m^2}{g^2}\right] - E\left[\prod_{j=1}^{i}\sqrt{\frac{G_j}{g}}\right]^2\right\}$$

$$= 4R^2 P_0 S_{sp} B_e (lg)^{2k-i} \cdot \left[\left(\frac{E[G^2]}{g^2}\right)^{k-i} - \left(\frac{E[\sqrt{G}]^2}{g}\right)^i\right]$$

171

Because $E[\sqrt{G}]^2 = E[G] - \text{Var}(\sqrt{G})$ and $\text{Var}\sqrt{G} \geq 0$, $E[\sqrt{G}]^2 \leq E[G] = g$. Therefore, $\left(\frac{E[\sqrt{G}]^2}{g}\right)^i \leq 1$. Hence,

$$\left(\sigma_I^2\right)_{sig-sp,i} \geq 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot (lg)^{k-i} \left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^{k-i} - 1\right]$$

The total intensity noise from signal beating with noise $\sigma_{IG_{sig-sp}}^2$ is,

$$\sigma_{IG_{sig-sp}}^2 \geq 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot \sum_{i=1}^{k} \left\{\left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^{k-i} - 1\right] (lg)^{k-i}\right\}$$

$$= 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot \sum_{j=0}^{k-1} \left\{\left[lg\left(1 + \frac{\sigma_G^2}{g^2}\right)\right]^j - (lg)^j\right\}$$

$$= \begin{cases} 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot \left\{\frac{g^2}{\sigma_G^2}\left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^k - 1\right] - k\right\} & \text{if } lg = 1 \\[3mm] 4R^2 P_0 S_{sp} B_e \cdot (lg)^k \cdot \left\{\frac{\left[lg\left(1 + \frac{\sigma_G^2}{g^2}\right)\right]^k - 1}{lg\left(1 + \frac{\sigma_G^2}{g^2}\right) - 1} - \frac{(lg)^k - 1}{lg - 1}\right\} & \text{otherwise} \end{cases} \qquad \square$$

## C.5   Derivation of $\sigma_{IG_{sp-sp}}^2$ in Eq. (5.31)

*Proof.* Define $\left(\sigma_I^2\right)_{sp-sp,ij}, (i < j)$ to be the intensity noise from the $i$th ASE term beating with the $j$th ASE term. The corresponding photocurrent is

$$I_{sp-sp,ij} = 2R \int\int (S_{sp} \cdot lG_{i+1} \cdot lG_{i+2} \cdots lG_k)^{\frac{1}{2}} \cdot (S_{sp} \cdot lG_{j+1} \cdot lG_{j+2} \cdots lG_k)^{\frac{1}{2}} \cos\theta_2 d\omega_n d\omega_n'$$

172

Therefore,

$$\left(\sigma_I^2\right)_{sp-sp,ij} = \text{Var}\left(I_{sp-sp,ij}\right)$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot \text{Var}\left(\left(\prod_{m=i+1}^{k} \frac{G_m lg}{g}\right)^{\frac{1}{2}} \left(\prod_{n=j+1}^{k} \frac{G_n lg}{g}\right)^{\frac{1}{2}}\right)$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot (lg)^{k-i}(lg)^{k-j} \cdot \text{Var}\left(\prod_{m=i+1}^{j} \left(\frac{G_m}{g}\right)^{\frac{1}{2}} \prod_{n=j+1}^{k} \frac{G_n}{g}\right)$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot (lg)^{k-i}(lg)^{k-j} \cdot \left\{ E\left[\left(\prod_{m=i+1}^{j} \left(\frac{G_m}{g}\right)^{\frac{1}{2}} \prod_{n=j+1}^{k} \frac{G_n}{g}\right)^2\right] - E\left[\prod_{m=i+1}^{j} \left(\frac{G_m}{g}\right)^{\frac{1}{2}} \prod_{n=j+1}^{k} \frac{G_n}{g}\right]^2 \right\}$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot (lg)^{k-i}(lg)^{k-j} \cdot \left\{ E\left[\prod_{n=j+1}^{k} \frac{G_n^2}{g^2}\right] - E\left[\prod_{m=i+1}^{j} \sqrt{\frac{G_m}{g}}\right]^2 \right\}$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot (lg)^{k-i}(lg)^{k-j} \cdot \left[\left(\frac{E[G^2]}{g^2}\right)^{k-j} - \left(\frac{E[\sqrt{G}]^2}{g}\right)^{j-i}\right]$$

$$\geq 4R^2 S_{sp}^2 B_o B_e \cdot (lg)^{k-i}(lg)^{k-j} \cdot \left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^{k-j} - 1\right]$$

The total intensity noise from ASE noise beating with itself is,

$$\sigma_{IG_{sp-sp}}^2 \geq \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left(\sigma_I^2\right)_{sp-sp,ij}$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot \sum_{i=1}^{k-1} (lg)^{k-i} \sum_{j=i+1}^{k} (lg)^{k-j} \left[\left(1 + \frac{\sigma_G^2}{g^2}\right)^{k-j} - 1\right]$$

$$= 4R^2 S_{sp}^2 B_o B_e \cdot \sum_{i=1}^{k-1} (lg)^{k-i} \sum_{j=i+1}^{k} \left[\left(lg\left(1 + \frac{\sigma_G^2}{g^2}\right)\right)^{k-j} - (lg)^{k-j}\right]$$

173

If $lg \neq 1$,

$$\sigma^2_{IG_{sp-sp}} = 4R^2 S^2_{sp} B_o B_e \cdot \sum_{i=1}^{k-1} (lg)^{k-i} \left\{ \frac{\left[ lg\left(1 + \frac{\sigma^2_G}{g^2}\right)\right]^{k-i} - 1}{lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1} - \frac{(lg)^{k-i} - 1}{lg - 1} \right\}$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \sum_{i=1}^{k-1} \left\{ \frac{\left[ (lg)^2 \left(1 + \frac{\sigma^2_G}{g^2}\right)\right]^{k-i} - (lg)^{k-i}}{lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1} - \frac{(lg)^{2(k-i)} - (lg)^{k-i}}{lg - 1} \right\}$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \sum_{i=1}^{k-1} \left\{ \frac{\left[ (lg)^2 \left(1 + \frac{\sigma^2_G}{g^2}\right)\right]^{k-i} - (lg)^{k-i}}{lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1} - \frac{(lg)^{2(k-i)}}{lg - 1} + \frac{(lg)^{k-i}}{lg - 1} \right\}$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \sum_{i=1}^{k-1} \left\{ \frac{\left[ (lg)^2 \left(1 + \frac{\sigma^2_G}{g^2}\right)\right]^{k-i}}{lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1} - \frac{(lg)^{2(k-i)}}{lg - 1} + \frac{lg\frac{\sigma^2_G}{g^2}(lg)^{k-i}}{\left[ lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1\right](lg - 1)} \right\}$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \left\{ \frac{(lg)^2 \left(1 + \frac{\sigma^2_G}{g^2}\right)}{lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1} \cdot \frac{\left[ (lg)^2 \left(1 + \frac{\sigma^2_G}{g^2}\right)\right]^{k-1} - 1}{(lg)^2 \left(1 + \frac{\sigma^2_G}{g^2}\right) - 1} - \frac{(lg)^2}{lg - 1} \cdot \frac{(lg)^{2(k-1)} - 1}{(lg)^2 - 1} + \frac{(lg)^2 \frac{\sigma^2_G}{g^2}\left[(lg)^{k-1} - 1\right]}{\left[ lg\left(1 + \frac{\sigma^2_G}{g^2}\right) - 1\right](lg - 1)^2} \right\}$$

If $lg = 1$,

$$\sigma^2_{IG_{sp-sp}} = 4R^2 S^2_{sp} B_o B_e \cdot \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left[ \left(1 + \frac{\sigma^2_G}{g^2}\right)^{k-j} - 1\right]$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \sum_{i=1}^{k-1} \left\{ \frac{\left(1 + \frac{\sigma^2_G}{g^2}\right)^{k-i} - 1}{\frac{\sigma^2_G}{g^2}} - (k - i) \right\}$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \sum_{n=1}^{k-1} \left\{ \frac{\left(1 + \frac{\sigma^2_G}{g^2}\right)^{n} - 1}{\frac{\sigma^2_G}{g^2}} - n \right\}$$

$$= 4R^2 S^2_{sp} B_o B_e \cdot \left\{ \frac{\left(1 + \frac{\sigma^2_G}{g^2}\right)\left[\left(1 + \frac{\sigma^2_G}{g^2}\right)^{k-1} - 1\right]}{\left(\frac{\sigma^2_G}{g^2}\right)^2} - \frac{k-1}{\frac{\sigma^2_G}{g^2}} - \frac{k(k-1)}{2} \right\}$$

□

174

# Bibliography

[1] Multiprotocol label switching. http://en.wikipedia.org/wiki/MPLS. [Online; accessed 01-May-2014].

[2] The zettabyte eratrends and analysis. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.html. [Online; accessed 01-May-2014].

[3] G. P. Agrawal. *Fiber-Optic Communication Systems*. Wiley-Interscience, Rochester, NY, third edition edition, 2002.

[4] S. Alexander, R. Bondurant, D. Byrne, V. Chan, S. Finn, R. Gallager, B. Glance, H. Haus, P. Humblet, R. Jain, I. Kaminow, M. Karol, R. Kennedy, A. Kirby, H. Le, A. Saleh, B. Schofield, J. Shapiro, N. Shankaranarayanan, R. Thomas, R. Williamson, and R. Wilson. A precompetitive consortium on wide-band all-optical networks. *Lightwave Technology, Journal of*, 11(5):714–735, May-Jun 1993.

[5] S. R. Asmussen. Steady-state properties of of gi/g/1. In *Applied Probability and Queues*, volume 51 of *Stochastic Modelling and Applied Probability*, pages 266–301. Springer New York, 2003.

[6] D. Banerjee and B. Mukherjee. A practical approach for routing and wavelength assignment in large wavelength-routed optical networks. *Selected Areas in Communications, IEEE Journal on*, 14(5):903 –908, jun 1996.

[7] R. Barry and P. Humblet. Models of blocking probability in all-optical networks with and without wavelength changers. *Selected Areas in Communications, IEEE Journal on*, 14(5):858–867, 1996.

[8] V. Chan. Optical access networks. In *Thirty-Ninth Annual Allerton Conference on Communication, Control, and Computing, (Invited)*, 2001.

[9] V. Chan. Optical flow switching networks. *Proceedings of the IEEE*, 100(5):1079 –1091, may 2012.

[10] V. Chan, A. Ganguly, and G. Weichenberg. Optical flow switching with time deadlines for high-performance applications. 2009.

[11] V. Chan, A. Kirby, and A. Saleh. AT&T/DEC/MIT precompetitive consortium on wideband all-optical networks. In *Lasers and Electro-Optics Society Annual Meeting, 1993. LEOS '93 Conference Proceedings. IEEE*, pages 4 –5, nov 1993.

[12] M. Chiani, D. Dardari, and M. K. Simon. New exponential bounds and approximations for the computation of error probability in fading channels. *Wireless Communications, IEEE Transactions on*, 2(4):840–845, July 2003.

[13] I. Chlamtac, A. Ganz, and G. Karmi. Lightnet: lightpath based solutions for wide bandwidth wans. In *INFOCOM '90, Ninth Annual Joint Conference of the IEEE Computer and Communication Societies. The Multiple Facets of Integration. Proceedings, IEEE*, pages 1014–1021 vol.3, Jun 1990.

[14] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: an approach to high bandwidth optical wan's. *Communications, IEEE Transactions on*, 40(7):1171 –1182, jul 1992.

[15] N. Christofides. *Graph Theory: An Algorithmic Approach (Computer Science and Applied Mathematics)*. Academic Press, Inc., Orlando, FL, USA, 1975.

[16] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.

[17] M. Gagnaire, M. Koubaa, and N. Puech. Network dimensioning under scheduled and random lightpath demands in all-optical wdm networks. *Selected Areas in Communications, IEEE Journal on*, 25(9):58–67, December 2007.

[18] M. R. Garey and D. S. Johnson. *Computers and Intractability-A Guide to the Theory of NP Completeness*. W. H. Freeman, San Franscisco, 1979.

[19] K. C. Guan, M. I. of Technology. Dept. of Electrical Engineering, and C. Science. *Cost-effective optical network architecture: a joint optimization of topology, switching, routing and wavelength assignment*. 2007.

[20] C. Henry. Theory of spontaneous emission noise in open resonators and its application to lasers and optical amplifiers. *Lightwave Technology, Journal of*, 4(3):288–297, Mar 1986.

[21] H. Huang and V. Chan. Transport layer protocol for optical flow-switched networks. In *International Conference on Communications (ICC 2013), 2013 IEEE*, June 2013.

[22] H. P. Huang, M. I. of Technology. Dept. of Electrical Engineering, and C. Science. *Transport Layer Protocol Design over Flow-Swithced Data Networks*. 2012.

[23] J. Junio, L. Zhang, and V. Chan. Design of long haul fast turn-on/off of lightpaths and induced physical layer characteristics of edfa-amplified meshed networks. In *OFC 2014. Optical Fiber Conference. Proceedings*, March 2014.

[24] R. Krishnaswamy and K. Sivarajan. Design of logical topologies: a linear formulation for wavelength-routed optical networks with no wavelength changers. *Networking, IEEE/ACM Transactions on*, 9(2):186–198, Apr 2001.

[25] M. Lax. Theory of laser noise. In *Proc. SPIE 1376*, volume 1376, March 1991.

[26] K. Lin, M. I. of Technology. Dept. of Electrical Engineering, and C. Science. *Green optical network design: power optimization of wide area and metropolitan area networks*. 2011.

[27] E. Mannie. Generalized multi-protocol label switching (gmpls) architecture. RFC 3945. IETF, Oct. 2004.

[28] S. Mookherjea and V. Chan. Remotely-pumped optical distribution networks. In *Lasers and Electro-Optics, 2001. CLEO '01. Technical Digest. Summaries of papers presented at the Conference on*, pages 497 –498, 2001.

[29] S. Pachnicke, E. Gottwald, P. Krummrich, and E. Voges. Combined impact of raman and edfa transients on long haul transmission system performance. In *Optical Communication (ECOC), 2007 33rd European Conference and Ehxibition of*, pages 1–2, Sept 2007.

[30] Y. Pan, D. Kilper, A. Morea, J. Junio, and V. Chan. Channel power excursions in gmpls end-to-end optical restoration with single-step wavelength tuning. In *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pages 1–3, March 2012.

[31] R. Ramaswami and K. Sivarajan. Routing and wavelength assignment in all-optical networks. *Networking, IEEE/ACM Transactions on*, 3(5):489 –500, oct 1995.

[32] R. Ramaswami and K. Sivarajan. Design of logical topologies for wavelength-routed optical networks. *Selected Areas in Communications, IEEE Journal on*, 14(5):840–851, Jun 1996.

[33] A. Schrijver. In *Combinatorial Optimization: Polyhedra and Efficiency*, volume 1 of *Algorithms and Combinatorics*, page 114. Springer, 2003.

[34] J. M. Simmons. *Optical network design and planning*. Springer, New York, 2008.

[35] Y. Sun, A. Srivastava, J. Zyskind, J. Sulhoff, C. Wolf, and R. Tkach. Fast power transients in wdm optical networks with cascaded edfas. *Electronics Letters*, 33(4):313–314, Feb 1997.

[36] C. Tian and S. Kinoshita. Analysis and control of transient dynamics of edfa pumped by 1480- and 980-nm lasers. *Lightwave Technology, Journal of*, 21(8):1728–1734, Aug 2003.

[37] H. L. V. Trees. *Detection, Estimation, and Modulation Theory Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. Krieger Publishing Co., Inc., Melbourne, FL, USA, 1992.

[38] S. M. V.W.S. Chan, S. Chan. Optical distribution networks. In *Optical Networks Magazine*, volume 3, pages 25–33, January/February 2002.

[39] G. Weichenberg, V. Chan, and M. Medard. Access network design for optical flow switching. In *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, pages 2390–2395, Nov. 2007.

[40] G. Weichenberg, V. Chan, and M. Medard. Design and analysis of optical flow-switched networks. *Optical Communications and Networking, IEEE/OSA Journal of*, 1(3):B81–B97, August 2009.

[41] G. Weichenberg, V. Chan, and M. Medard. Performance analysis of optical flow switching. In *Communications, 2009. ICC '09. IEEE International Conference on*, pages 1–6, June 2009.

[42] G. E. Weichenberg, M. I. of Technology. Dept. of Electrical Engineering, and C. Science. *Design and analysis of optical flow switched networks*. 2009.

[43] A. Yariv. *Quantum Electronics*. John Wiley & Sons, third edition edition, 1989.

[44] E. Yetginer, Z. Liu, and G. Rouskas. Fast exact ilp decompositions for ring rwa. *Optical Communications and Networking, IEEE/OSA Journal of*, 3(7):577–586, July 2011.

[45] S. Yoo, W. Xin, L. Garratt, J. Young, G. Ellinas, J. C. Chiao, M. Rauch, J. Baran, B. Meagher, H. Leblanc, and G.-K. Chang. Observation of prolonged power transients in a reconfigurable multiwavelength network and their suppression by gain-clamping of optical amplifiers. *Photonics Technology Letters, IEEE*, 10(11):1659–1661, Nov 1998.

[46] H. Zang and J. P. Jue. A review of routing and wavelength assignment approaches for wavelength-routed optical wdm networks. *Optical Networks Magazine*, 1:47–60, 2000.

[47] L. Zhang and V. Chan. Scalable fast scheduling for optical flow switching using sampled entropy and mutual information broadcast. *J. Opt. Commun. Netw.*, 6(5):459–475, May 2014.

[48] X. Zheng, L. Zhang, and V. Chan. Metropolitan and access network architecture design for optical flow switching (submitted). In *Global Telecommunications Conference, 2014. GLOBECOM '14. IEEE*, pages 1–6, Dec. 2014.