DOCTORAL THESIS OF SORBONNE UNIVERSITY

Scientific Field

Information and Communication Sciences and Technologies

Doctoral School: EDITE ED 130

Computer Science, Telecommunications, and Electronics of Paris

Presented by

Mr. Chenyu SUN

To obtain the degree of

Doctor of Philosophy

Thesis Title:

Digital Twins, AI and Autonomous Optical Networks

Defended on 2nd March 2026 before the jury composed of

Ms. Carmen MAS MACHUCA Professor, University of the Bundeswehr Munich, Germany	Reviewer
Mr. Daniel KILPER Professor, Trinity College Dublin, Ireland	Reviewer
Mr. Adlen KSENTINI - *Jury President* Professor, EURECOM, France	Examiner
Mr. Konstantinos CHRISTODOULOPOULOS Assistant Professor, University of Athens, Greece	Examiner
Mr. Petros ELIA Professor, EURECOM, France	Advisor
Mr. Photios A. STAVROU Assistant Professor, EURECOM, France	Co-supervisor
Mr. Yvan POINTURIER Researcher, Huawei Technologies France, France	Co-supervisor

THÈSE DE DOCTORAT DE SORBONNE UNIVERSITÉ

Spécialité

Sciences et Technologies de l'Information et de la Communication

École Doctorale : EDITE ED 130

Informatique, Télécommunications et Électronique de Paris

Présentée par

M. Chenyu SUN

Pour obtenir le grade de

Doctorat

Sujet de la thèse :

Jumeaux Numériques, IA et Réseaux Optiques Autonomes

Soutenue le 2 Mars 2026 devant le jury composé de

Mme. Carmen MAS MACHUCA Professeure, Université de la Bundeswehr à Munich, Allemagne	Rapporteure
M. Daniel KILPER Professeur, Trinity College Dublin, Irlande	Rapporteur
M. Adlen KSENTINI - *Président du Jury* Professeur, EURECOM, France	Examinateur
M. Konstantinos CHRISTODOULOPOULOS Maître de Conférence, Université d'Athènes, Grèce	Examinateur
M. Petros ELIA Professeur, EURECOM, France	Directeur
M. Photios A. STAVROU Maître de Conférence, EURECOM, France	Co-encadrant
M. Yvan POINTURIER Chercheur, Huawei Technologies France, France	Co-encadrant

À ce quart de ma vie passé à Paris,
la brise éternelle de la rivière,
et la lune lumineuse.

"If you are lucky enough to have lived in Paris as a young man, then wherever you go for the rest of your life, it stays with you, for Paris is a moveable feast."

– E. Hemingway, *A Moveable Feast* (1964)

八聲甘州暮雨忘機
六州歌頭弦外知音
七賢會飲蜉蝣間世
九渡舟虛乘物游心

"One must speak very much before one keeps silent."

– FUNG Yu-Lan, *A Short History of Chinese Philosophy* (1948)

Acknowledgements

I would like to express my deepest gratitude to my supervisors for their guidance and support throughout my doctoral studies. My industrial supervisor from Huawei Paris Research Center, Yvan, has an exceptional ability to identify emerging technologies and promising research directions. With his many years of experience and profound understanding of optical networks, he has offered invaluable insights. I am truly grateful for the opportunity to explore the future of optical networks alongside such a distinguished expert in the field. My academic supervisor from EURECOM, Photis, possesses remarkable expertise in mathematics, algorithms, and AI. His inspiration has extended far beyond the scope of my research topics; our discussions have broadened my horizons well beyond optical communications.

I am also sincerely thankful to all my colleagues in the Optical Communication Technology Lab of Huawei Paris Research Center. This lab was established in 2019, the same year that I came to France. I decided then that I would pursue a CIFRE PhD in optical communications here. In 2022, I joined the lab for my Master's final year internship, and in 2023, I began my doctoral journey. Gabriel, the director of the laboratory, is an exceptionally intelligent researcher with over two decades of experience in optical communications. I am grateful to him for creating this research lab, from which I have learned a lot. My internship supervisor, Alessio, is also highly talented. During the half year we worked together, I gained a deep understanding of modern software-controlled architectures for optical networks, which laid the foundation for my subsequent PhD work. My heartfelt thanks go to Xin and Reda from the optical networking team; without their support, many of my achievements would not have been possible. I also thank our colleagues from the other teams – Abel, Dylan, Flavio, Hien, Loig, Manuel, Mihai, Nayla, Salma, Shuqi, Tarek, Tug, Xiaohui, and many others. They are all kind and wonderful people, and working with them and sharing experiences has been truly enjoyable. I greatly appreciate all the help and support they have provided, and I truly cherish the many memorable moments we shared both at conferences and in everyday life.

Besides, I am grateful for the inspiring discussions with peers at academic conferences and for the fruitful collaboration with the researchers from PoliMi.

In 2024, I participated in the BFC program at INSEAD Business School, which opened the door to a new world for me. I am grateful for the memorable time I spent with my classmates and professors during that program.

Since arriving in France in 2019, I have also been fortunate to meet many friends and classmates at SupOp, Paris-Saclay, IP Paris, EURECOM and Sorbonne over the past seven years. Their companionship has greatly enriched my experience.

Finally, I would like to thank my family and long-time friends. Even though distance and time often kept us apart, their constant support has always reminded me that I never walk alone.

"Water has no exciting flavour, but one never grows tired of it."

Chenyu SUN
Spring 2026
Paris, France

Abstract

Modern optical transport networks must ensure error-free transmission under dynamic conditions and heterogeneous ROADM-based infrastructures. Maintaining high quality of transmission (QoT) is challenging due to interdependent launch-power settings across optical multiplex sections, where naïve adjustments may induce transient signal-to-noise ratio (SNR) degradation. Increasing system complexity further motivates the integration of physical-layer modeling with intelligent automation.

This thesis presents proof-of-concept contributions toward autonomous optical network control and management by combining digital twin (DT) technology with AI-based approaches.

First, DT-enabled multi-step lookahead strategies are proposed for autonomous power equalization. With monitoring data, the DT predicts QoT before applying configuration changes. By evaluating sequences of control actions with fixed and adaptive step sizes, DT-enabled closed-loop control mitigates QoT degradation during intermediate steps and avoids suboptimal convergence, while DT-assisted parallel configuration improves efficiency. Experiments on a commercial ROADM-based optical network testbed demonstrate fast and safe network reconfiguration without intermediate SNR margin degradation.

Second, an AI-agent-based framework is developed for intent-driven control and management. A control agent translates high-level intents into APIs calls for automated configuration, while a management agent leverages retrieval-augmented generation to analyze alarms and support root-cause analysis. Beyond task automation, the proposed agent architecture illustrates how AI agents can coordinate decision-making, integrate domain knowledge, and support the evolution toward autonomous optical network operation.

Overall, this work demonstrates the feasibility of combining DT-based closed-loop control with AI-agent-based automation, providing a foundation for future autonomous optical networks.

Résumé

Les réseaux de transport optiques modernes doivent garantir une transmission sans erreur dans des conditions dynamiques et des infrastructures ROADM hétérogènes. Le maintien d'une haute qualité de transmission (QoT) est complexe en raison de l'interdépendance des puissances de lancement entre sections de multiplexage optique, où des ajustements naïfs peuvent dégrader le rapport signal sur bruit (SNR). La complexité croissante des systèmes motive l'intégration de modèles de couche physique avec des mécanismes d'automatisation intelligents.

Cette thèse propose des contributions exploratoires pour le contrôle et la gestion autonomes des réseaux optiques, en combinant jumeaux numériques (DT) et intelligence artificielle (IA).

Premièrement, des stratégies multi-étapes avec anticipation, basées sur DT, sont proposées pour l'égalisation autonome des puissances. À partir des données de supervision, le DT prédit la QoT avant application des configurations. L'évaluation de séquences d'actions avec pas fixes et adaptatifs permet un contrôle en boucle fermée limitant la dégradation de la QoT aux étapes intermédiaires et évitant les convergences sous-optimales, tandis que la configuration parallèle améliore l'efficacité. Des expériences sur une plateforme commerciale démontrent une reconfiguration rapide et sûre, sans dégradation intermédiaire des marges de SNR.

Deuxièmement, un cadre basé sur des agents d'IA est proposé pour un contrôle et une gestion pilotés par intentions. Un agent de contrôle traduit les intentions en appels d'API pour l'automatisation de la configuration, tandis qu'un agent de gestion exploite la génération augmentée par récupération pour analyser les alarmes et assister l'identification des causes racines. Cette architecture illustre le rôle des agents d'IA dans la coordination des décisions et l'intégration des connaissances métier.

Ces travaux démontrent la faisabilité d'un contrôle en boucle fermée basé sur DT combiné à une automatisation par agents d'IA, et posent les bases des futurs réseaux optiques autonomes.

Contents

List of Figures

List of Tables

Acronyms and Abbreviations

AI	Artificial intelligence
ADC	Analog to digital converter
AWG	Arrayed waveguide grating
B2B	Back-to-back
BA	Booster amplifier
BER	Bit error rate
CD	Chromatic dispersion
DAC	Digital to analog converter
DSP	Digital signal processing
DT	Digital twin
E2E	End-to-end
EDFA	Erbium-doped fiber amplifier
EGN model	Enhanced Gaussian noise model
FEC	Forward error correction
FWM	Four-wave mixing
GN model	Gaussian noise model
GSNR	Generalized signal-to-noise ratio
HCF	Hollow core fiber
ILA	In-line amplifier
IM-DD	Intensity modulation-direct detection
ITU	International telecommunication union
LEAF	Large effective area fiber
LLM	Large language model
ML	Machine learning
MZM	Mach–Zehnder modulator
NE	Network element

NF	Noise figure
NLI	Nonlinear interference
NLP	Natural language processing
NLSE	Nonlinear Schrödinger equation
NSR	Noise-to-signal ratio
OA	Optical amplifier
OADM	Optical add/drop multiplexer
OCh	Optical channel
OCM	Optical channel monitor
OMS	Optical multiplex section
OOK	On–off keying
OPM	Optical performance monitor
OSA	Optical spectrum analyzer
OSNR	Optical signal-to-noise ratio
OTN	Optical transport network
OTS	Optical transport section
OXC	Optical cross connect
P2P	Point-to-point
PA	Pre-amplifier
PAM	Pulse-amplitude modulation
PBS	Polarizing beam splitter
PCS	Probabilistic constellation shaping
PDL	Polarization dependent loss
PDM	Polarization division multiplexing
PSCF	Pure silica core fiber
PSK	Phase shift keying
QAM	Quadrature amplitude modulation
QoT	Quality of transmission
QPSK	Quadrature phase shift keying
ROADM	Reconfigurable optical add/drop multiplexer
RX	Receiver
SDN	Software defined networking
SHB	Spectrum hole burning
SMF	Single mode fiber

SNR	Signal-to-noise ratio
SOA	Semiconductor optical amplifier
SOP	State of polarization
SPM	Self-phase modulation
SRS	Stimulated Raman scattering
SSFM	Split-step Fourier method
SSMF	Standard single mode fiber
TRX	Transceiver
TX	Transmitter
VOA	Variable optical attenuator
WDL	Wavelength dependent loss
WDM	Wavelength division multiplexing
WSS	Wavelength selective switch
XPM	Cross-phase modulation

Chapter 1

Introduction

1.1 Optical Communications and Optical Networks

Optical communication systems constitute the physical foundation of modern backbone, metro and data center networks, enabling high speed, long-reach, and highly reliable data transport. In such networks, services are carried by light paths, i.e., modulated optical signals assigned to a specific physical route and a spectral "slot" within the wavelength grid. The light signal propagates transparently from transmitter to receiver through optical fibers, amplifiers, filters, and multiplexers, forming the end-to-end optical channel (OCh) or service. Multiple optical channels are transmitted in a common fiber with wavelength division multiplexing (WDM) technique, sharing the same physical medium and therefore interacting through linear and nonlinear effects.

To ensure reliable service delivery, operators require that transmitted signals remain error-free after forward error correction (FEC) decoding, typically targeting post-FEC bit-error rate (BER) below 10^{-15}. However, as optical signals propagate, they accumulate distortions and noise that degrade their physical-layer quality of transmission (QoT). The main impairments include: wavlength-dependent loss (WDL); amplified spontaneous emission (ASE) noise generated by optical amplifiers; nonlinear Kerr effects in optical fibers; filtering penalties; polarization-dependent loss (PDL); and receiver noise, including implementation penalties associated with digital signal processing (DSP) blocks. In extended-band systems—such as 6 THz C-band operation or multi-band C+L systems—stimulated Raman scattering (SRS) introduces power transfer from shorter to longer wavelengths, further affecting channel performance.

Although most of these impairments are well understood and described by physical models, some components exhibit behavior that remains difficult to predict

accurately. In particular, erbium-doped fiber amplifiers (EDFAs), the dominant amplifier technology in current optical networks, are known to exhibit input-load-dependent and wavelength-dependent, i.e. spectrum hole burning (SHB) effect. These effects lead to inaccuracies in predicting amplifier gain profiles and output spectra using purely physics-based models. As a result, researchers have explored machine learning (ML) techniques to improve model accuracy and compensate for incomplete or imperfect parameter knowledge [1–6]. Similar challenges also arise in other components whose behavior is only partially characterized or where the model inputs (e.g., span loss, connector losses, or exact gain tilt) are uncertain.

To mitigate these inaccuracies, numerous monitoring-assisted or data-driven techniques have been proposed [5, 7, 8]. In practice, model inaccuracies translate into the need for larger design and operational margins to guarantee error-free transmission under all conditions [9, 10]. These margins reduce spectral efficiency and limit network throughput. Lowering margins is possible but generally increases operational complexity: for example, during rapid service restoration following a failure, operators may temporarily assume a reduced margin to accelerate light-path provisioning, at the risk of temporarily sub-optimal QoT [11].

The increasing complexity of coherent transceivers, elastic spectrum allocation, multi-band operation, and disaggregated optical networks has motivated the community toward autonomous optical networks, analogous to self-driving systems in other engineering domains. Such systems leverage the rich telemetry made available by advanced optical performance monitors (OPMs) and coherent receivers, using this real-time data to adaptively optimize the network. Autonomous optical networks are expected to:

- continuously re-optimize launch power, routing, and spectrum assignments when excess margin is detected;
- detect, classify, and mitigate physical-layer degradations before service outages occur;
- dynamically adjust modulation formats or symbol rates to maintain service availability;
- reconfigure resources proactively to isolate failing components or circumvent fiber degradation.

1.2 Challenges

A central component of this vision is the digital twin (DT), a software representation of the physical network that uses real-time monitoring data to estimate QoT, emu-

late commissioning strategies, forecast performance under hypothetical actions, and determine safe reconfiguration commands before applying them to the live network. DTs have already been demonstrated for power equalization, capacity maximization, and resilience improvement [12, 13]. Re-optimizing optical networks through per-channel launch power adjustment, commonly referred to as power equalization—has received significant attention in both academia and industry [14–20].

However, existing approaches typically assume that only the final target power profile matters: the DT computes the optimal per-channel launch power for each OMS, and this configuration is subsequently pushed to the network in a single shot. In reality, intermediate reconfiguration steps matter: changing the launch power of one or more channels perturbs the QoT of all co-propagating channels via nonlinear interference, power coupling, amplifier gain dynamics, and SRS [21–23]. As a result, sequentially adjusting channels without considering transient states may temporarily degrade some light paths below their FEC threshold, potentially causing outages.

In an autonomous operational setting, this behavior is unacceptable: all existing services must remain above the FEC limit throughout the entire re-optimization process. Therefore, one of the key challenges addressed in this thesis is how to determine a safe sequence of reconfiguration steps—possibly with restrictions on step size and ordering—that transitions the network from its current state to the target optimized state, while keeping all services error-free at every intermediate step?

DT inaccuracies must also be considered. Even with advanced modeling, a DT may mismatch the physical network due to unknown parameters, partial monitoring, or model limitations. Therefore, during reconfiguration, conservative margins and step-by-step actuation are required to ensure robustness: if the DT detects unexpected QoT degradation, it must update its internal representation and adjust subsequent steps accordingly.

Once safety is guaranteed, a second major challenge concerns speed. A naïve baseline applies one reconfiguration action at a time, fully sequentially. However, many actions satisfy a form of independence: adjusting powers on two OMSs that do not share services may be performed in parallel without affecting each other. More generally: Some power changes have minimal impact on other channels and can be parallelized; Filter reconfiguration times scale sub-linearly with the number of channels adjusted simultaneously; Amplifier settling times and WSS reconfiguration delays may differ and must be coordinated. This introduces a combinatorial scheduling problem, where the goal is to minimize total reconfiguration time while

respecting QoT safety constraints.

In addition to the physical-layer challenges discussed above, another fundamental challenge in the operation of optical networks arises from the increasing complexity of network management tasks and the growing risk of human-induced errors. Modern optical transport systems involve tens of OMSs, hundreds of services, and thousands of parameters related to routing, spectrum allocation, amplifier settings, alarms, and operational procedures. Even highly trained engineers inevitably face difficulties when dealing with repetitive configuration tasks, large volumes of logs, time-critical restoration procedures, or non-standard network states. Human mistakes—such as misconfigured channels, incorrect power settings, improper wavelength assignment, or delayed reaction to alarms—remain one of the leading sources of service disruption in operational networks. As networks continue to scale in capacity and heterogeneity, the cognitive load required to manually operate them becomes increasingly unsustainable.

More recently, the rapid progress of artificial intelligence (AI) has further expanded the scope of data-driven approaches in optical networks. Beyond conventional ML models, emerging techniques such as deep learning, large language models (LLMs), and AI agents offer new capabilities for handling complex, multi-layer decision-making and network automation. In particular, AI agents can leverage domain knowledge, telemetry data, and high-level objectives to assist in network control and orchestration, complementing traditional model-based and optimization-driven methods.

While DT-enabled power optimization focuses on maintaining QoT safety during physical-layer reconfiguration, an LLM-based agent can automate a wide spectrum of control-plane tasks: translating high-level operator intents into valid configuration commands, preventing syntactic or semantic errors in network operation, analyzing complex multi-source alarm logs, and assisting during troubleshooting or restoration. These tasks are traditionally prone to human error, especially when performed under time pressure or in large-scale networks with heterogeneous equipment and vendor-specific interfaces. This motivates the use of AI, and in particular LLM-based AI agents, as a complementary automation layer that addresses challenges beyond power equalization.

However, deploying LLM-based AI agents in optical networks introduces its own challenges. The agent must reliably transform natural-language instructions into correct API calls, respecting network topology constraints, resource availability, and safety rules. Any misinterpretation can lead to misconfiguration of services, spectrum conflicts, or accidental service interruption. Furthermore, the agent must

integrate seamlessly with the DT to validate the physical feasibility of any requested action, ensuring that the LLM does not inadvertently propose reconfigurations that violate QoT constraints. Unlike power optimization, which is grounded in numerical models, network automation tasks require combining symbolic reasoning (topology, configuration, alarms) with physical awareness (QoT, FEC margins), making the interaction between the LLM and the DT essential.

Therefore, while AI agents holds significant promise for reducing human mistakes and enabling fully autonomous optical network operation, its integration requires careful design of interfaces, fine-tuning with domain-specific data, strict validation of generated commands, and coordination with DT simulations. Addressing these challenges is a necessary step toward achieving safe, reliable, and operator-trusted AI-driven optical network automation.

1.3 Research Questions and Contributions

The increasing complexity of optical transport networks, driven by WDM technologies, flexible ROADMs, and even disaggregated network architectures, has made network operation and optimization increasingly challenging. Optical networks involve many interacting physical-layer effects, such as nonlinear interference, amplifier dynamics, and filtering penalties, which make manual configuration and optimization difficult and error-prone. At the same time, the availability of large volumes of monitoring data and the emergence of AI techniques open new opportunities for autonomous and intelligent network operation.

In this context, this thesis investigates how DT and AI can be used to enable safe, efficient, and autonomous operation of optical networks. The work is structured around the following research questions (RQ).

RQ1: How can autonomous launch-power equalization be safely performed in operational optical networks without causing transient QoT degradation during intermediate reconfiguration steps?

While power optimization or equalization has been widely studied, many existing approaches focus only on the final optimized configuration. However, in real operational networks, intermediate reconfiguration steps can temporarily degrade QoT due to nonlinear interference, power transfer, and amplifier dynamics. Ensuring that all services remain above their FEC thresholds during the entire reconfiguration process is therefore a critical challenge.

To address this problem, this thesis proposes autonomous power equalization strategies based on multi-step lookahead prediction, dynamic step-size adjustment,

and safe parallel configuration validated through DTs. These techniques enable the network to transition safely from a non-optimized state to an optimized configuration while maintaining service continuity.

The proposed algorithms are experimentally validated on a heterogeneous optical network testbed composed of commercial ROADMs, EDFAs, WSSs, and coherent transponders, demonstrating improvements in SNR margin and significant reductions in commissioning time. This work is mainly presented in Chapter 4.

Publications related to RQ1:

1. **C. Sun** et al., "Digital twin-enabled multi-step strategies for autonomous power equalization in optical networks," *Journal of Optical Communications and Networking (JOCN)*, vol. 17, no. 7, C41-C50, 2025. **[ECOC2024 Top-Scored Invited]** [24]
2. **C. Sun** et al., "Digital Twin-Enabled Optical Network Channel Power Management by WSS and Booster Auto-Adjustment," *Optical Fiber Communications Conference and Exhibition (OFC)*, 30 March - 3 April 2025, San Francisco, CA, USA. **[Poster]** [25]
3. **C. Sun** et al., "Digital Twin Enabled Automatic Power Adjustment with Multi-Step Lookahead Prediction," *50th European Conference on Optical Communication (ECOC)*, 22-26 September 2024, Frankfurt, Germany. **[Oral, Top Scored - Upgraded to Invited]** [26]
4. **C. Sun** et al., "Digital Twin-Enabled Optical Network Automation: Power Re-Optimization," *Optical Fiber Communications Conference and Exhibition (OFC)*, 24-28 March 2024, San Diego, CA, USA. **[Oral]** [27]
5. **C. Sun** et al., "Digital Twin-Enabled Service Optimization Sequence of Actions for Power Equalization," *Asia Communications and Photonics Conference and The International Photonics and OptoElectronics Meetings (ACP /POEM)*, 4-7 November 2023, Wuhan, China. **[Oral, Best Paper Award in Industry Innovation]** [28]

RQ2: How can AI, and in particular LLMs, assist the automation and lifecycle management of optical networks?

In addition to physical-layer optimization, operating optical networks requires handling complex configuration procedures, analyzing large volumes of alarm logs, and even interacting with heterogeneous vendor-specific interfaces. These tasks are often repetitive and prone to human errors. Recent advances in LLMs provide new opportunities to improve automation and human–machine interaction in network management systems.

This thesis explores the integration of locally deployed and fine-tuned LLM-

based AI agents with a DT-enabled network controller framework. The proposed system enables intent-based control through automatic API generation, automated alarm-log analysis, documentation querying through retrieval-augmented generation, and multilingual interaction for network operation.

These capabilities are experimentally demonstrated on the commercial testbed, showing the feasibility of LLM-assisted lifecycle automation for optical networks. This work is mainly presented in Chapter 5.

Publications related to RQ2:

1. **C. Sun** et al., "Experimental demonstration of local AI-Agents for lifecycle management and control automation of optical networks," *Journal of Optical Communications and Networking (JOCN)*, vol. 17, no. 8, C82-C92, 2025. **[ECOC2024 PDP Invited]** [29]
2. **C. Sun** et al., "First experimental demonstration of full lifecycle automation of optical network through fine-tuned LLM and digital twin," *50th European Conference on Optical Communication (ECOC)*, 22-26 September 2024, Frankfurt, Germany. **[Post-Deadline Paper, PDP]** [30]
3. **C. Sun** et al., "Demonstration of LLM-based AI-Agent for Optical Network Management and Automation," *50th European Conference on Optical Communication (ECOC)*, 22-26 September 2024, Frankfurt, Germany. **[Demo]** [31]

Overall, the contributions of this thesis demonstrate the feasibility of combining DT and AI to enable safe and efficient automation of optical networks. The results provide initial experimental validation of DT-enabled power control strategies and AI agents-assisted network automation, representing an important step toward the realization of autonomous optical networks.

Beyond these proof-of-concept studies, this work also advances the transition from research to industrial applications by introducing localized AI agent frameworks tailored for operator environments. These developments have been demonstrated in collaboration with major network operators, including China Mobile [32] and Etisalat-UAE [33], highlighting the practical viability and deployment potential of the proposed approaches.

1.4 Thesis Outline

The remainder of this thesis is organized as follows:

Chapter 2 introduces the physical foundations of optical fibers and optical communications, from electromagnetic wave propagation and fiber properties to coher-

ent transceivers, optical amplifiers, and ROADM-based network architectures.

Chapter 3 presents the physical-layer noise sources, QoT metrics, the GN/EGN models, and power optimization and power control principles that underpin the proposed DT framework.

Chapter 4 details the architecture of the DT, the methodology for safe autonomous power equalization, and the experimental validation on realistic optical network topologies, addressing Research Question 1.

Chapter 5 explores the use of AI and large language models for optical network control, including the design of local LLM-based AI agents and their integration with the DT to achieve full lifecycle automation, addressing Research Question 2.

Finally, Chapter 6 concludes the thesis, summarizes the main findings, and outlines promising directions for future research.

Chapter 2

From an Optical Fiber to Optical Networks

This chapter starts with physical optics, the principles of optical fiber communications and introduce the architecture of optical transmission systems and networks.

2.1 Light: Wave and Particle

Light exhibits a dual nature, behaving both as a wave and as a particle, a concept central to modern physics and photonics. This duality allows us to describe and exploit light in different ways depending on the phenomenon of interest.

2.1.1 Light as an Electromagnetic Wave

In the wave description, light is an electromagnetic wave, consisting of oscillating electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields that are perpendicular to each other and to the direction of propagation. The behavior of these fields is rigorously described by *Maxwell's equations* [34]:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{2.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \tag{2.4}$$

where $\nabla\cdot$ is the divergence operator, $\nabla\times$ is the curl operator, ρ is the density of electrical charges, $\mathbf{J}$ is the electric current density vector, ε_0 is the permittivity of the free space μ_0 is the permeability of the free space.

In free space or in the non-conductive core of an optical fiber ($\rho = 0$, $\mathbf{J} = 0$), these equations reduce to the *electromagnetic wave equations* [35]:

$$\nabla^2\mathbf{E} - \mu_0\varepsilon_0\frac{\partial^2\mathbf{E}}{\partial t^2} = 0, \tag{2.5}$$

$$\nabla^2\mathbf{B} - \mu_0\varepsilon_0\frac{\partial^2\mathbf{B}}{\partial t^2} = 0. \tag{2.6}$$

The solutions to these equations are typically expressed as plane waves:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0\, e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.7}$$

$$\mathbf{B}(\mathbf{r},t) = \mathbf{B}_0\, e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.8}$$

where $\mathbf{k}$ is the wave vector, $\mathbf{r}$ is the position vector, ω is the angular frequency, t is the time, and $\mathbf{E}_0$, $\mathbf{B}_0$ are the field amplitudes.

The phase velocity of light in vacuum is given by [35]:

$$v_p = \frac{\omega}{|\mathbf{k}|} = \frac{1}{\sqrt{\mu_0\varepsilon_0}} = c, \tag{2.9}$$

and in a medium with refractive index n, the phase velocity is reduced [35]:

$$v_p = \frac{c}{n}. \tag{2.10}$$

This wave-based description provides the foundation for understanding mode propagation, dispersion, and interference in optical fibers.

Having described light as an electromagnetic wave with oscillating electric and magnetic fields, it becomes important to consider the orientation of these oscillations. In most practical optical systems, the direction of the electric field vector is of particular interest, as it determines how the wave interacts with materials, optical components, and the fiber itself. This property, known as polarization, characterizes the trajectory traced by the tip of the electric field vector over time and provides essential insight into phenomena such as birefringence, polarization mode dispersion, and the operation of polarization-sensitive devices.

For a plane wave (Eq. (2.7)) propagating along the z-axis, the electric field vector can be decomposed into two polarization waves x-pol and y-pol as:

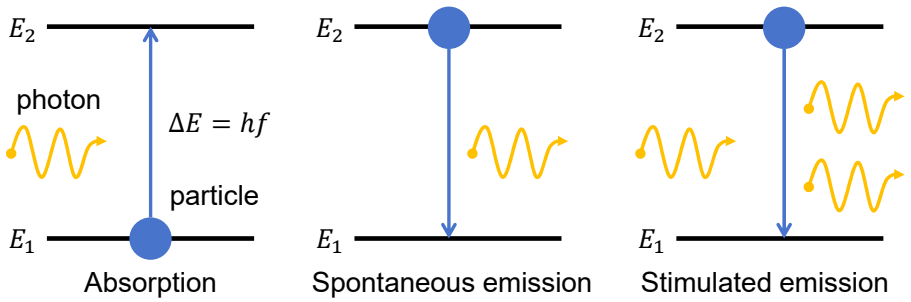


Figure 2.1: Phenomena of photons. E_1 and E_2 are the two different energy levels of the particle, ΔE is the energy gap between them.

$$\mathbf{E}(x, y, z, t) = \frac{E_0}{\sqrt{2}} e^{i(kz-\omega t)} \hat{\mathbf{x}} + \frac{E_0}{\sqrt{2}} e^{i(kz-\omega t+\phi)} \hat{\mathbf{y}}, \tag{2.11}$$

where ϕ is the phase difference between the components x-pol and y-pol, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors in the x and y directions.

2.1.2 Light as a Particle (Photon)

In the particle description, light consists of discrete packets of energy called *photons*. Each photon carries energy proportional to its frequency, described by Planck's relation [36]:

$$E = hf = h\frac{c}{\lambda}, \tag{2.12}$$

where E is the photon energy, h is Planck's constant (6.626×10^{-34} J $\cdot$ s), c is the light velocity in vacuum, f is the frequency and λ is the wavelength. The particle nature of light explains phenomena such as the photoelectric effect, spontaneous and stimulated emission (shown in the Fig. 2.1), and the operation of lasers and optical amplifiers.

In optical communications, the photon perspective is particularly important for understanding:

- Amplification in devices such as erbium-doped fiber amplifiers (EDFA),
- Quantum-limited noise in detectors,
- Nonlinear interactions where photon-photon effects influence signal integrity.

2.2 Optical Fibers

According to different principles, light can propagate through glass (silica) [37] or vacuum media. This section will introduce the principles of light propagation in

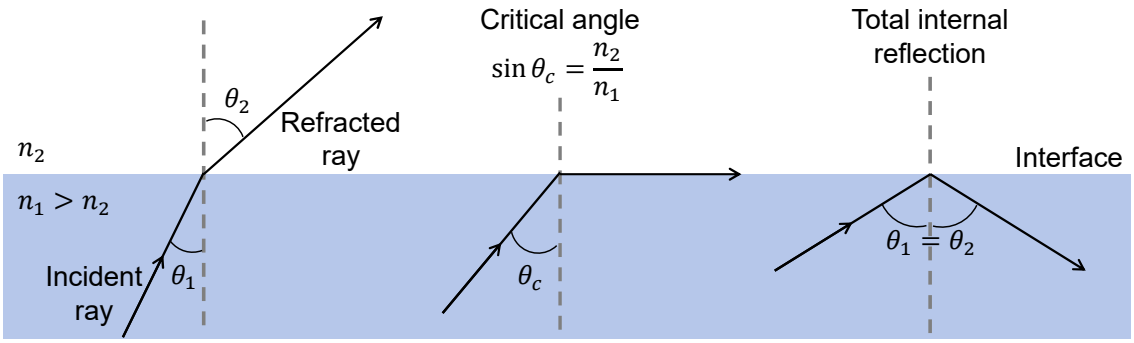


Figure 2.2: Snell's law and total internal reflection of light.

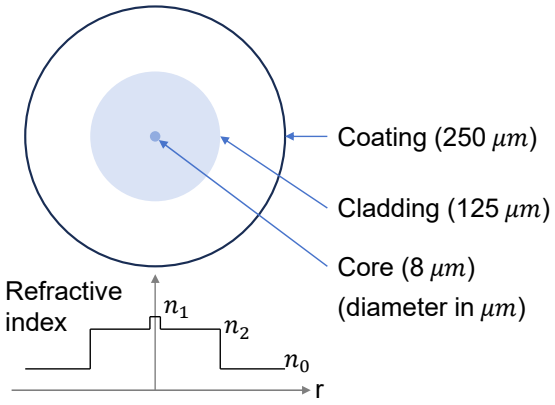


Figure 2.3: Scheme of ITU G.652 standard single mode fiber (SSMF).

various optical fibers, focusing on how light travels through traditional silica fibers and the physical laws it follows.

2.2.1 Traditional Silica Fibers

Total Internal Reflection

Traditional silica optical fibers confine light waves within the core through the principle of total internal reflection. Snell's Law (also known as the law of refraction) describes how light bends when it passes from one medium to another with a different refractive index. It is mathematically expressed as [38]:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2, \tag{2.13}$$

where n_1 and n_2 are the refractive indices of the first and second media, respectively, θ_1 is the angle of incidence (the angle between the incident ray and the normal as shown in Fig. 2.2), θ_2 is the angle of refraction (the angle between the refracted ray and the normal as shown in Fig. 2.2).

By designing the refractive index of the core to be greater than that of the

cladding, light waves can propagate from one end of the fiber to the other without leakage. Fig. 2.3 shows the cross section and refractive index profile of standard single mode fiber (ITU G.652).

Light Propagation and Nonlinear Schrödinger Equation

Once confined within the fiber core by total internal reflection, the optical field propagates along the longitudinal axis and experiences a variety of physical effects that influence its amplitude, phase, and polarization. The evolution of the slowly varying complex envelope $A(z,t)$ of the optical field in a silica fiber can be accurately described by the *nonlinear Schrödinger equation* (NLSE) [39], which accounts for fiber loss, chromatic dispersion, and Kerr nonlinearity:

$$\frac{\partial A(z,t)}{\partial z} = -\frac{\alpha}{2}A(z,t) - i\frac{\beta_2}{2}\frac{\partial^2 A(z,t)}{\partial t^2} + i\gamma|A(z,t)|^2 A(z,t), \tag{2.14}$$

where:

- $A(z,t)$: slowly varying envelope of the optical field,
- z: propagation distance,
- t: retarded time in the frame moving at the group velocity,
- α: fiber attenuation coefficient,
- β_2: group velocity dispersion parameter,
- $\gamma = \frac{2\pi}{\lambda}\frac{n_{\mathrm{NL}}}{A_{\mathrm{eff}}}$: nonlinear coefficient,
- n_{NL}: nonlinear refractive index,
- A_{eff}: effective area.

Fiber Loss

The attenuation coefficient $\alpha(\lambda)$ represents the wavelength-dependent loss (WDL) of optical power due to absorption and scattering in the optical fiber material. The transmitted optical power decreases exponentially with distance as follows:

$$P(z) = P(0)e^{-\alpha z}. \tag{2.15}$$

Typical attenuation coefficients for SSMF are around 0.2 dB/km near 1550 nm as shown in Fig. 2.4, mainly limited by Rayleigh scattering and intrinsic material absorption.

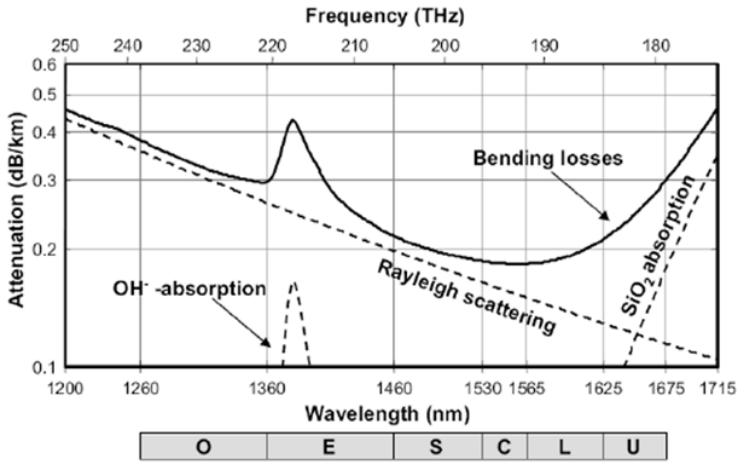


Figure 2.4: WDL profile of silica fiber and transmission bands.

Chromatic Dispersion

Dispersion arises from the wavelength dependence of the propagation constant $\beta(\lambda)$. The second-order term β_2 accounts for group velocity dispersion (GVD), which causes different spectral components of a pulse to travel at different group velocities, leading to pulse broadening. When $\beta_2 > 0$ (normal dispersion), higher frequencies travel slower; when $\beta_2 < 0$ (anomalous dispersion), higher frequencies travel faster.

The dispersion coefficient D is used to characterize the fiber dispersion [39]:

$$D = -\frac{2\pi c}{\lambda^2}\beta_2, \tag{2.16}$$

the typical dispersion coefficient value D for SSMF is approximately 17 ps/nm/km at 1550 nm.

Dispersion plays a crucial role in determining the transmission distance and bandwidth of high-speed systems. It can be compensated using dispersion compensating fibers, fiber Bragg gratings, or digital signal processing (DSP) in coherent receivers.

Kerr Nonlinearity

The nonlinear term $i\gamma|A|^2A$ represents the Kerr effect, where the refractive index depends on the instantaneous optical intensity $I = |A|^2$. This induces self-phase modulation (SPM), which broadens the optical spectrum, cross-phase modulation (XPM) between WDM channels, and four-wave mixing (FWM) among different frequency components. These effects are particularly significant in long-haul and dense WDM systems, where high optical powers and long propagation distances

accumulate nonlinear phase shifts.

The interplay between dispersion and nonlinearity governs many nonlinear optical phenomena, such as soliton propagation in the anomalous dispersion regime or modulation instability. The NLSE thus serves as the fundamental model for understanding and simulating signal evolution in optical transmission systems.

Polarization Effects and the Manakov Equation

In real SSMFs, the optical field has two orthogonal polarization components that experience slightly different propagation constants due to birefringence. When birefringence varies rapidly along the fiber, the polarization evolution can be statistically averaged, leading to the *Manakov equation* [40]:

$$\frac{\partial \mathbf{A}}{\partial z} = -\frac{\alpha}{2}\mathbf{A} - i\frac{\beta_2}{2}\frac{\partial^2 \mathbf{A}}{\partial t^2} + i\frac{8}{9}\gamma\|\mathbf{A}\|^2\mathbf{A}, \tag{2.17}$$

where $\mathbf{A} = [A_x(z,t), A_y(z,t)]^T$ is the Jones vector of the two polarization components. The Manakov equation effectively describes nonlinear propagation on polarization average and is widely used in coherent optical communication system modeling, particularly for multiplexed signals of polarization-division.

The NLSE and its vectorial extensions, such as the Manakov equation, form the theoretical foundation for analyzing optical signal propagation in fibers. They capture the essential physical mechanisms—loss, dispersion, and nonlinearity—that dictate the performance limits and design strategies of modern high-capacity WDM and coherent transmission systems.

2.2.2 Hollow Core Fibers

Hollow-core fibers represent a paradigm shift in optical transmission. Instead of confining light within a solid silica core, HCFs guide light through an air or gas-filled core, typically using photonic bandgap or anti-resonant guiding mechanisms. Since most of the optical power propagates in air, nonlinear effects and latency are significantly reduced compared to solid-core fibers.

Recent advances in anti-resonant hollow-core fibers (AR-HCFs) [41] have allowed record-low attenuation values (<0.1dB/km) [42] and broad transmission windows across multiple bands. These fibers exhibit group velocities closer to the speed of light in vacuum, reducing latency by up to 30% relative to silica fibers.

Despite these advantages, HCFs still face challenges related to fabrication complexity, splicing losses, gas line absorption, and mode coupling stability, but con-

tinued improvements in preform design and draw techniques suggest promising prospects for integration into future high-capacity optical transport networks.

2.3 Modulation and Coherent Optical Transceivers

From Eq. (2.11), it can be observed that we can modulate light waves through four degrees of freedom: amplitude, phase, wavelength (or frequency) and polarization. This section will introduce the principles of optical wave modulation, i.e., how we load the information we want to communicate onto light waves, and the working principles of coherent optical transceivers.

2.3.1 Modulation Formats

The modulation format defines how information is encoded onto an optical carrier for transmission through a fiber channel.

In amplitude modulation, information is carried by varying the optical power of the signal. The simplest form, known as on–off keying (OOK), represents binary data using the presence ("1") or absence ("0") of light shown in Fig. 2.5(a). Despite its simplicity and ease of implementation with direct detection, OOK is limited

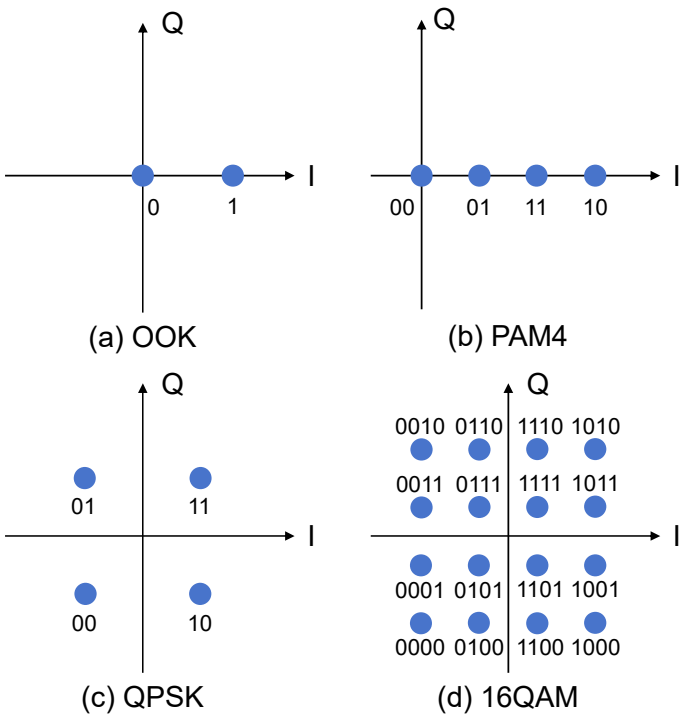


Figure 2.5: Constellation diagrams of different modulation formats. (I: in-phase, Q: quadrature)

in spectral efficiency and sensitivity to noise. To increase the number of bits per symbol, multi-level modulation formats such as pulse amplitude modulation (PAM) are widely adopted. Each symbol corresponds to one of M discrete amplitude levels, providing a spectral efficiency $SE = \log_2(M)$ bits/s/Hz under Nyquist signaling. For instance, PAM-4 encodes two bits per symbol (Fig. 2.5(b)), effectively doubling the data rate compared to OOK without increasing the symbol rate.

Phase modulation, on the other hand, encodes information in the instantaneous phase of the optical carrier. In binary phase-shift keying (BPSK), two symbols are represented by phase shifts of 0 and π radians. Quadrature phase-shift keying (QPSK) extends this principle to four equally spaced phase states separated by $\pi/2$, thereby transmitting two bits per symbol, as shown in Fig. 2.5(c). By jointly modulating both amplitude and phase, two-dimensional constellations are formed, giving rise to quadrature amplitude modulation (QAM). In M-QAM, each symbol is mapped to a unique point in the complex plane, corresponding to one of M combinations of amplitude and phase. For instance, the constellation diagram of 16QAM is shown in Fig. 2.5(d). In a conventional M-ary quadrature amplitude modulation (M-QAM), all constellation symbols are transmitted with equal probability, resulting in a fixed number of bits per symbol given by $b = \log_2(M)$.

However, symbols located at the edges of the constellation carry higher energy and are more vulnerable to noises, which will be introduce in Section 3.1. Probabilistic constellation shaping (PCS) improves transmission efficiency by assigning non-uniform probabilities to symbols [43], favoring those closer to the origin of the constellation diagram. This reduces the average symbol energy and enables operation closer to the Shannon capacity limit.

The average information carried by each symbol can be expressed by the entropy:

$$H(X) = -\sum_{i=1}^{M} P(x_i) \log_2 P(x_i), \tag{2.18}$$

where $P(x_i)$ is the probability of transmitting the i-th constellation point x_i.

For uniform M-QAM, $P(x_i) = 1/M$ and $H(X) = \log_2(M)$, the spectrum efficiency is:

$$SE = \log_2(M) \text{ bits/s/Hz}. \tag{2.19}$$

In PCS, the symbol probabilities are intentionally made non-uniform—typically following a Maxwell–Boltzmann distribution—to increase the occurrence of low-energy symbols. E.g., a PCS-16QAM constellation shape is shown in the Fig. 2.6. This reduces the average symbol energy and improves the achievable information rate for a given signal-to-noise ratio (SNR). Consequently, the entropy $H(X)$ be-

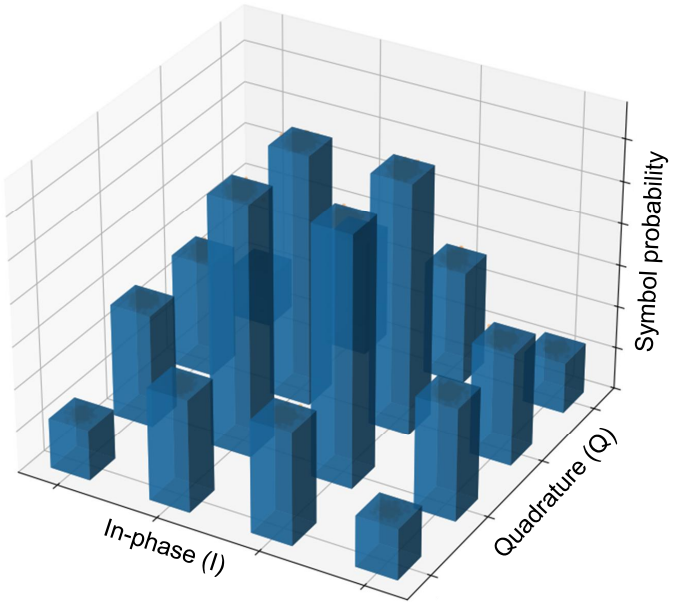


Figure 2.6: Symbol probability constellation of PCS-16QAM.

comes smaller than $\log_2(M)$, meaning that the average number of information bits per symbol is reduced, then the spectrum efficiency is:

$$SE = H(X) \text{ bits/s/Hz}. \tag{2.20}$$

Another degree of freedom in optical communication is polarization. Since an optical wave can be decomposed into two orthogonal polarization components, it is possible to transmit two independent data streams simultaneously on these components—a technique known as polarization-division multiplexing (PDM). This effectively doubles the transmission capacity without increasing the required optical bandwidth. In coherent optical systems, PDM is typically combined with complex modulation formats such as QPSK or QAM, leading to configurations like PDM-QPSK or PDM-16QAM.

Beyond single-carrier modulation, the wavelength of light provides an additional dimension for multiplexing. Wavelength-division multiplexing (WDM) enables multiple optical carriers, each at a distinct wavelength, to be modulated independently and transmitted simultaneously over the same fiber. Each wavelength serves as a separate communication channel, dramatically increasing total system capacity.

The total capacity of a PDM-WDM system can be expressed as:

$$C = 2\sum_{k}^{N_\lambda} B(k) \cdot SE(k) \cdot (1 - OH_{\text{FEC}}(k)), \tag{2.21}$$

where 2 is the factor of two polarization components, N_λ is the number of channels and $B(k)$ is the bandwidth of channel-k, $SE(k)$ is the spectrum efficiency of channel-k, $OH_{\text{FEC}}(k)$ is the overhead of FEC coding.

In summary, the choice of modulation format in optical communication systems reflects a trade-off between implementation complexity, spectral efficiency, and system robustness. By exploiting the four physical degrees of freedom—amplitude, phase, polarization, and wavelength—and incorporating advanced techniques such as probabilistic shaping, modern coherent transceivers achieve transmission rates approaching the fundamental capacity limits of optical fiber channels.

2.3.2 Transmitter

The transmitter in a coherent optical transceiver is responsible for converting the digital data into an optical signal that can be transmitted over long distances. The scheme is shown as Fig. 2.7. The process begins with the data source, which generates the digital data to be transmitted. This data can come from various sources, such as routers or switches in a communication network.

After the data is generated, it undergoes FEC encoding. FEC adds redundancy to the data, which helps in detecting and correcting errors that may occur during

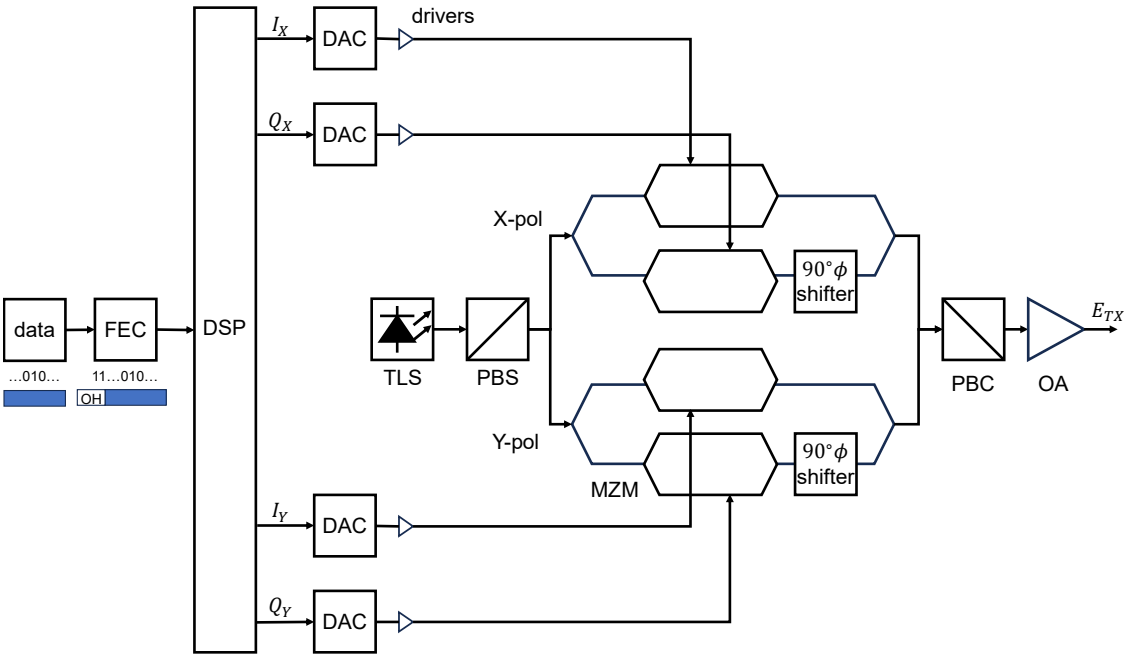


Figure 2.7: Coherent optical transmitter. TLS: tunable laser source; PBS: polarization beam splitter; PBC: polarization beam combiner; FEC: forward error correction encoding; OH: overhead of FEC; DSP: digital signal processing; DAC: digital to analog converter; MZM: Mach-Zehnder modulator; OA: optical amplifier.

transmission. This is crucial for maintaining the integrity of the data over long distances.

The encoded data is then processed by a digital dignal processor. The DSP generates the in-phase (I) and quadrature (Q) components of the signal. These components are essential for modulating the optical carrier. The I and Q components are typically generated using a digital-to-analog converter (DAC).

The next step involves modulating the laser beam with the I and Q components. A common modulator used in coherent systems is the Mach-Zehnder modulator (MZM). The MZM can modulate both the phase and amplitude of the optical signal, allowing for advanced modulation formats such as QPSK or 16QAM.

The modulated signal is then amplified using an optical amplifier. This amplification is necessary to boost the signal strength before it is transmitted through the optical fiber.

2.3.3 Receiver

The receiver in a coherent optical transceiver is responsible for converting the received optical signal back into the original digital data. The scheme is shown as Fig. 2.8.

The process begins with the reception of the optical signal from the optical fiber. A local oscillator (LO) laser generates a coherent reference signal. This reference

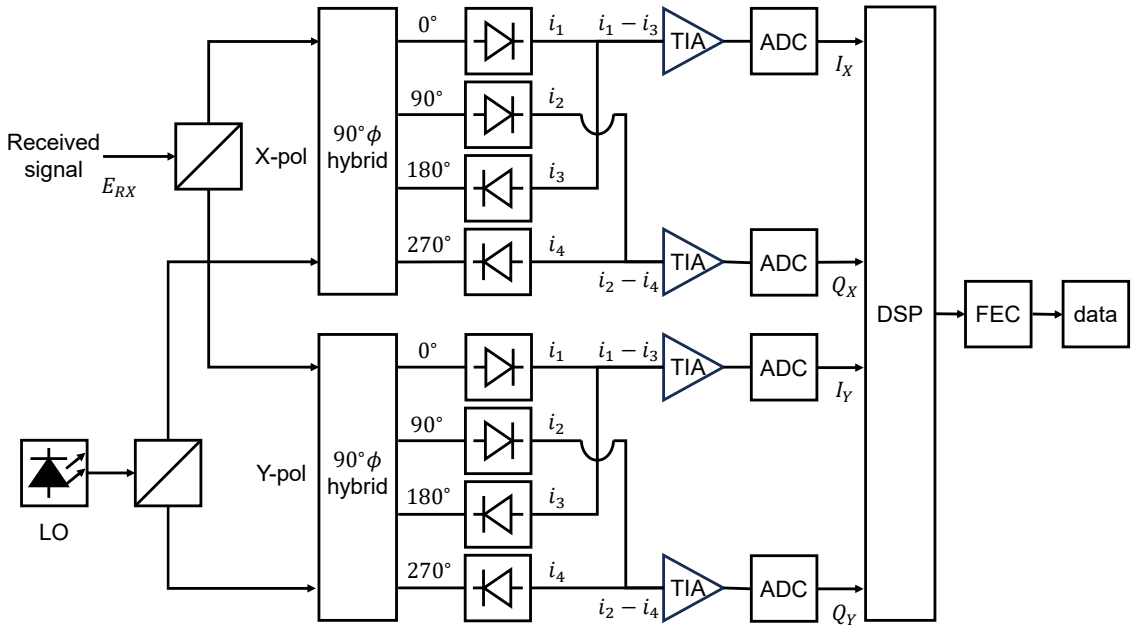


Figure 2.8: Coherent optical receiver. LO: local oscillator; TIA: Transimpedance amplifier; ADC: analog to digital converter; FEC: forward error correction decoding.

signal is crucial for the coherent detection process. The local oscillator signal is mixed with the received signal using an optical hybrid. The optical hybrid combines the received signal and the LO signal to produce four outputs: the in-phase (I) and quadrature (Q) components of the received signal and their conjugates.

These outputs are detected by photodetectors (PDs) and trans-impedance amplifiers (PIAs), which convert the optical signals into electrical signals. The electrical signals are then digitized using analog-to-digital converters (ADCs). The digitized signals are processed by DSP to recover the original data. The DSP performs various tasks, including equalization to correct for signal distortions, carrier recovery to synchronize the phase of the received signal with the local oscillator, and clock recovery to synchronize the timing of the received data.

Finally, the recovered data is decoded using FEC to correct any errors that may have occurred during transmission. The decoded data is then passed to the data sink, such as a router or switch, for further processing.

2.4 Optical Amplifiers

Optical amplifiers are essential components in modern optical communication systems, enabling the compensation of transmission losses over long fiber spans without the need for optical-electrical-optical (OEO) conversion. They amplify optical signals directly in the optical domain, thus improving system scalability and transparency across modulation formats and bit rates. Among the different types of optical amplifiers, rare-earth-doped fiber amplifiers and semiconductor optical amplifiers (SOAs) are the most widely deployed in WDM transmission systems.

2.4.1 Rare-Earth-Doped Fiber Amplifier

Rare-earth-doped fiber amplifiers rely on optical gain achieved through stimulated emission in a fiber core doped with trivalent rare-earth ions such as erbium (Er^{3+}), ytterbium (Yb^{3+}), or thulium (Tm^{3+}). When the doped fiber is pumped with light of an appropriate wavelength (typically from a laser diode), electrons in the dopant ions are excited to higher energy states. The population inversion thus created allows an incoming signal photon to stimulate emission of identical photons, resulting in optical amplification along the fiber length.

The Erbium-Doped Fiber Amplifier (EDFA) is the most widely used optical amplifier in long-haul and metro optical networks, primarily because its gain spectrum coincides with the C-band — region where silica fibers exhibit minimum attenua-

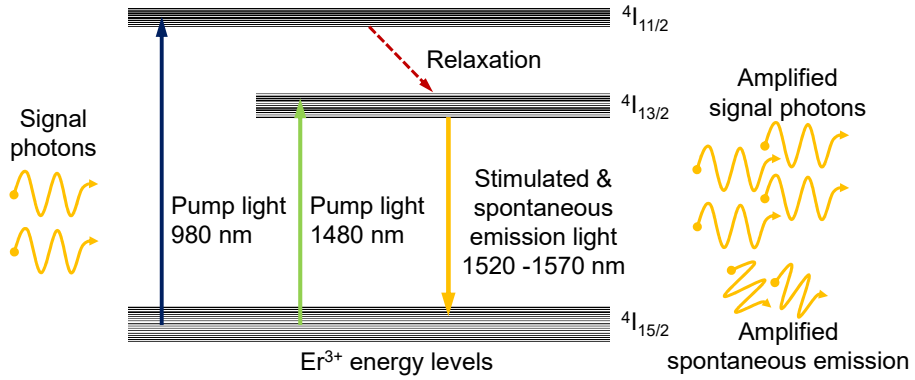


Figure 2.9: Erbium energy level scheme.

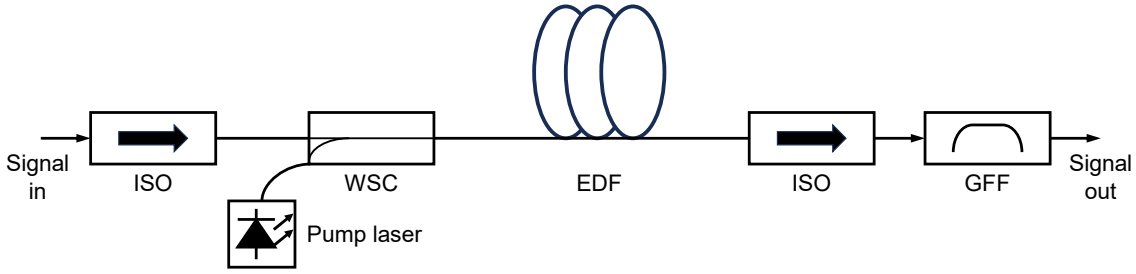


Figure 2.10: Basic scheme of an EDFA. EDF: erbium-doped fiber; ISO: isolator; WSC: wavelength selective coupler; GFF: gain flattening filter.

tion (as shown in Fig. 2.4) [44]. When the 980 nm or 1480 nm pump light excites the erbium ions to a higher metastable state ($^4I_{13/2}$ in Fig. 2.9), they can relax via stimulated emission to the ground state ($^4I_{15/2}$ in Fig. 2.9) when stimulated by incoming signal photons around 1550 nm, thus amplifying the signal. However, there is also a spontaneous phenomenon introduced in Section 2.1.2, which produces amplified spontaneous emission (ASE) noise.

As shown in Fig. 2.10, the EDFA operates by using an erbium-doped fiber (EDF) to amplify optical signals directly within the fiber. The process begins when a weak optical signal is input into the EDFA. An optical isolator is used at both the input and the output to prevent any backward-reflected light from causing instability. A pump laser, typically operating at 980 nm or 1480 nm, injects high-intensity light into the EDF via a wavelength selective coupler (WSC). This pump light excites the erbium ions to a higher energy state. As the signal light travels through the fiber, it stimulates the excited erbium ions to emit photons at the same wavelength as the signal, a process known as stimulated emission. This results in amplification of the signal. The amplified signal is then separated from the residual pump light using another WSC and output from the EDFA. Additionally, a gain flattening filters (GFF) can be used to ensure a more uniform gain across the entire signal bandwidth, which is crucial for maintaining signal quality in multi-channel systems.

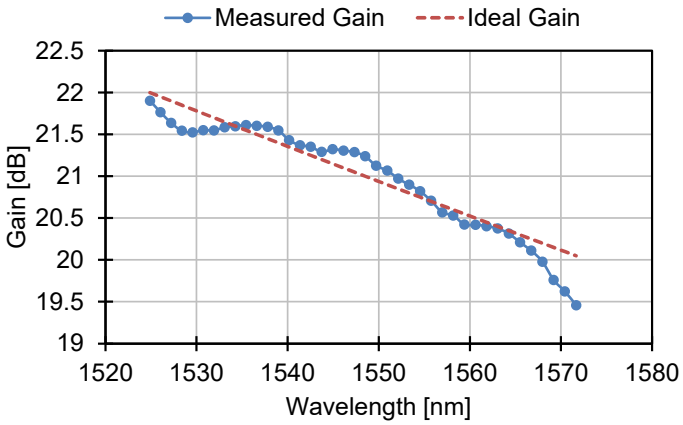


Figure 2.11: Gain profile of a commercial C-band EDFA working at 21dB gain-lock mode with 2dB tilt.

Two dominant operating modes are commonly considered:

- **Power-lock mode:** The total output power is maintained constant, typically via gain saturation or pump control. This introduces strong coupling between channels, as an increase in one channel's power must be compensated by a decrease in others.
- **Gain-lock mode:** The amplifier maintains a fixed gain for the total input power. Additionally, to compensate for the WDL of the fiber, the EDFA does not use a flat gain spectrum but instead sets a certain value of tilt, as shown in Fig. 2.11. While reducing direct coupling, wavelength-dependent gain ripple and spectral effects still induce inter-channel interactions.

The steady-state behavior of constant power amplifiers shows that channel power depends on the relative input powers of all channels, leading to inherent coupling across wavelengths [22]. Moreover, amplifier transient responses, governed by the population inversion dynamics, introduce time-dependent effects that propagate along cascaded spans.

2.4.2 Semiconductor Optical Amplifier

The SOA operates by utilizing the principles of stimulated emission in a semiconductor material, typically made of materials like indium phosphide (InP) or gallium arsenide (GaAs) [45]. When an optical signal passes through the active region of the SOA, it interacts with the excited electrons, causing them to drop to a lower energy state and emit additional photons, thereby amplifying the signal. SOAs are compact, integrable, and can be tuned over a wide range of wavelengths, making them highly versatile for various applications such as optical fiber communications, optical interconnects, and optical signal processing. Their ability to amplify signals

without the need for conversion to electrical form makes them crucial components in high-speed and long-distance communication networks.

Compared to EDFA, SOA offers several distinct advantages and disadvantages. SOAs are generally more compact and can be easily integrated with other semiconductor devices, making them ideal for on-chip applications and optical interconnects. They also have a broader gain bandwidth, which can be advantageous in WDM systems. However, SOAs typically have higher noise figures and are more susceptible to nonlinear effects, such as four-wave mixing and cross-gain modulation, which can degrade signal quality in high-power applications. In addition, SOAs have a lower saturation output power compared to EDFAs, which are known for their high gain and low noise characteristics, making EDFAs more suitable for long-haul optical communication systems. Despite these limitations, the flexibility and integration capabilities of SOAs make them valuable components in a variety of optical systems.

2.4.3 Raman Amplifier

Raman amplification is a distributed optical amplification technique based on the nonlinear process of stimulated Raman scattering (SRS) in silica fibers [46]. When a high-power pump wave co-propagates or counter-propagates with the signal, a portion of the pump energy is transferred to the signal through interactions with optical phonons. This energy transfer occurs when the frequency difference between pump and signal matches the characteristic Raman shift of silica, approximately 13.2 THz, thereby providing optical gain at the desired signal wavelength.

In contrast to discrete rare-earth-doped amplifiers such as EDFAs, Raman amplifiers do not require specially doped gain media. Instead, the transmission fiber itself acts as the gain medium, enabling the amplification to be distributed along the span rather than localized at specific points. Depending on the system design, Raman pumping may be launched in the forward, backward, or bidirectional direction, allowing flexible control of the gain profile and noise performance.

Raman amplifiers offer several advantages in modern optical communication systems [47]. First, the distributed nature of the gain significantly improves the noise figure of the span, often outperforming EDFAs in long-haul coherent links. Second, the Raman gain spectrum is highly flexible: by combining multiple pump wavelengths, it is possible to engineer broad and flat gain profiles that support wide WDM bandwidths. Finally, the improved signal power evolution along the fiber reduces nonlinear impairments and enhances the achievable transmission reach.

In practice, Raman amplifiers are widely deployed in long-haul and ultra-long-

haul coherent systems, typically in hybrid Raman–EDFA configurations. They are essential for extending optical bandwidth beyond the conventional C-band, improving OSNR margins, and supporting high-capacity WDM transmission. As optical networks continue to scale in capacity and reach, Raman amplification remains a key enabling technology for advanced optical transmission systems.

2.4.4 Noise Figure of Optical Amplifiers

The Noise Figure (NF) of an optical amplifier is a critical parameter that quantifies the degradation of the SNR due to the amplifier [48]:

$$\mathrm{NF} = \frac{\mathrm{SNR}_{\mathrm{in}}}{\mathrm{SNR}_{\mathrm{out}}}, \tag{2.22}$$

where SNR_{in} and SNR_{out} are the SNR in linear scale at the input and output of the OA.

In dB scale, it can be expressed as:

$$\mathrm{NF}_{\mathrm{dB}} = 10 \log_{10} \left(\frac{\mathrm{SNR}_{\mathrm{in}}}{\mathrm{SNR}_{\mathrm{out}}} \right). \tag{2.23}$$

The typical NF for different types of OAs can vary, with EDFAs generally having a NF of 4-6 dB, Raman amplifiers around 3-5 dB, and SOAs around 6-10 dB. A lower NF is generally preferred to maintain high quality of transmission in optical communication systems.

2.5 Wavelength Selective Switch

The use of wavelength selective switches (WSS) is integral to the construction and operation of optical cross-connects (OXCs), which are essential components in modern optical communication networks. OXCs enable dynamic and flexible routing of optical signals between multiple input and output ports, facilitating efficient management of network traffic and bandwidth. The integration of WSS in OXCs significantly enhances their functionality and performance by providing precise control over the wavelengths of light.

The working principle of WSS is based on the ability to selectively route individual wavelengths of light, as shown in Fig. 2.12. When a multi-wavelength optical signal, consisting of multiple channels, enters the WSS, it is first collimated and directed to a dispersive element, such as a diffraction grating or a prism. This dispersive element separates the different wavelengths into distinct beams, each

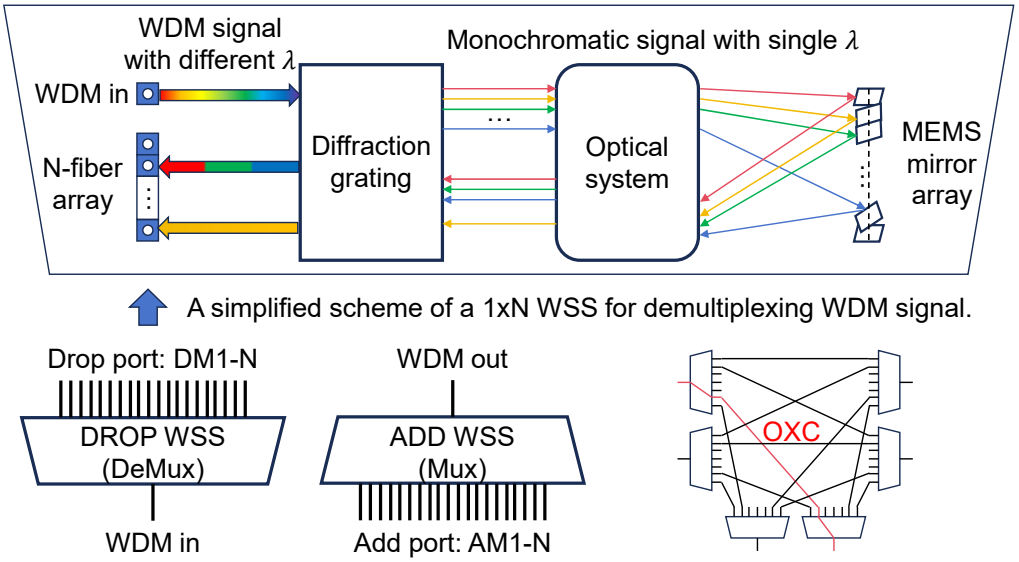


Figure 2.12: A simplified scheme of WSSs for adding (multiplexing, Mux) / dropping (demultiplexing, DeMux) WDM signal and building OXCs.

corresponding to a specific channel. The separated beams are then focused onto a tunable filter, which can be programmed to reflect or transmit each wavelength as needed, such as liquid crystal on silicon (LCoS) or micro-electron-mechanical systems (MEMS). This precise control over the wavelengths enables the WSS to route each channel to a specific output fiber, providing the flexibility needed for dynamic network reconfiguration.

Once the wavelengths are selectively reflected or transmitted by the tunable filter, they are recombined using another dispersive element, which combines the individual beams back into a multi-wavelength signal. This recombined signal is then focused and coupled to the output fiber, which can be connected to different nodes or devices in the network. The ability to dynamically route and switch wavelengths without affecting the other channels is crucial to managing network traffic, optimizing bandwidth utilization, and supporting various network architectures, such as ring, mesh, and point-to-point configurations.

In the context of OXCs, the integration of WSS technology provides several key benefits. First, it enables dynamic wavelength routing, allowing any input wavelength to be connected to any output port without interference. This non-blocking architecture ensures that the network can handle a high volume of traffic and complex routing scenarios without congestion. Second, WSS supports scalability as it can handle a large number of wavelengths and ports, making it suitable for growing networks. The ability to add or remove wavelengths without disrupting the existing network infrastructure ensures that the OXC can adapt to changing network

requirements.

Moreover, WSS devices are designed to minimize insertion loss and maximize isolation between different wavelengths. The low insertion loss ensures that signal strength is maintained throughout the network, reducing the need for additional amplification. High isolation prevents crosstalk between channels, ensuring that the quality of the optical signals is preserved. These characteristics are critical to maintaining the integrity of the data transmitted and to ensuring high network reliability and performance.

2.6 WDM Transmission Systems and Networking

2.6.1 Point-to-Point System

A WDM-based optical transport network consists of multiple optical multiplex sections (OMSs). The OMS is a critical component of optical transport networks, specifically designed to manage the adding/dropping of multiple wavelengths over a single fiber.

As shown in Fig. 2.13, in an OMS, a multiplexer (MUX), e.g. WSS, combines multiple optical signals, each at a different wavelength, into a single composite signal that can be transmitted over a single fiber. At the receiving end, a demultiplexer (DEMUX), e.g. WSS, separates the composite signal back into its individual wavelengths. This process allows for efficient use of fiber capacity, enabling the simultaneous transmission of multiple data streams.

The OMS also incorporates OAs to boost the signal strength over long distances, ensuring reliable transmission. As shown in Fig. 2.13, there are N+1 OAs for an OMS has N fiber spans. We refer to the first OA of each OMS as the booster

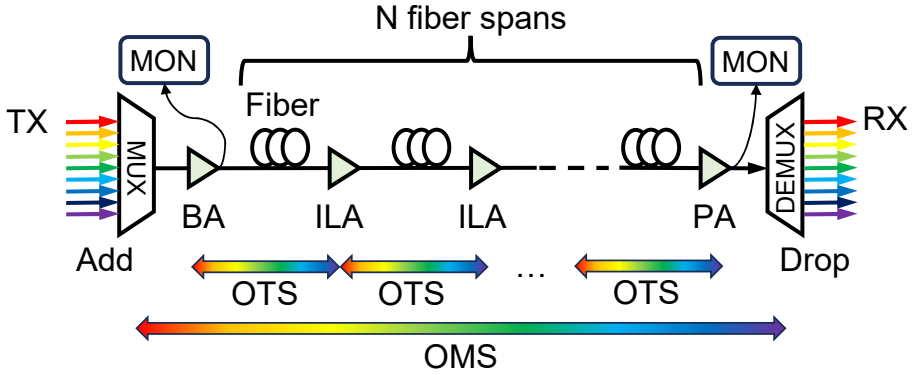


Figure 2.13: Scheme of an optical multiplex section. MUX: multiplexer; OTS: optical transmission section; BA: booster amplifier; ILA: in-line amplifier; PA: pre-amplifier; DEMUX: demultiplxer; MON: monitor.

amplifier (BA), the other OAs after fiber spans are called in-line amplifiers (ILAs), and the last OA is also called pre-amplifier (PA).

Additionally, the OMS supports various management and monitoring functions, which are essential for maintaining the integrity and performance of the optical network. For instance, by using the optical channel monitor (OCM) or optical performance monitor (OPM) at the output of BA, the launch power profile of the BA in an OMS can be tuned by adjusting the WSS attenuation profile, so that power equalization can be implemented to optimize the performance of services. Overall, the OMS plays a vital role in enhancing the capacity and flexibility of modern optical communication systems.

2.6.2 Optical Transport Network

Early WDM transmission systems were designed as fixed point-to-point optical links. In such systems, each wavelength was statically assigned between two endpoints, and any service reconfiguration required manual fiber reconnection or reengineering of intermediate fixed optical add/drop multiplexer (FOADM) nodes. These architectures lacked flexibility: wavelengths could not be rerouted, added, or removed without affecting the entire OMS. As network capacity and dynamic traffic patterns increased, this rigid structure became insufficient for modern transport requirements.

For large-scale optical networks, the reconfigurable optical add/drop multiplexers (ROADMs) fundamentally changed the topology of optical networks by enabling per-wavelength switching directly in the optical domain. The OMS provides the

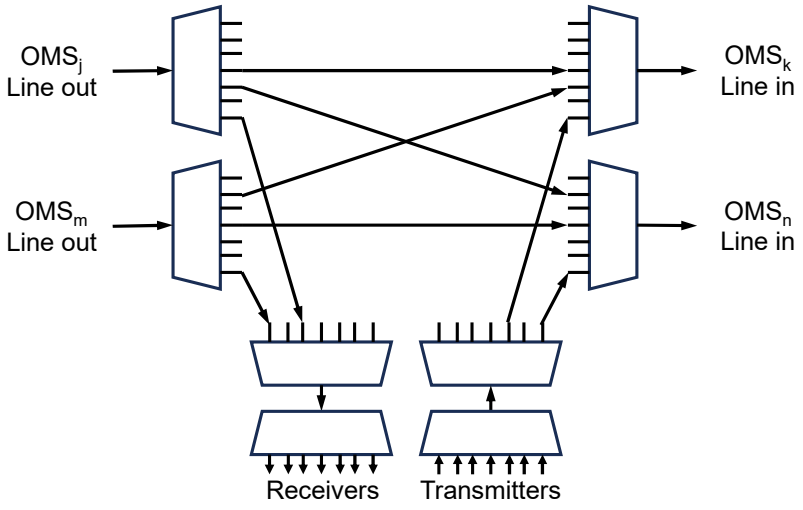


Figure 2.14: An example of reconfigurable optical add/drop multiplexer (ROADM) node.

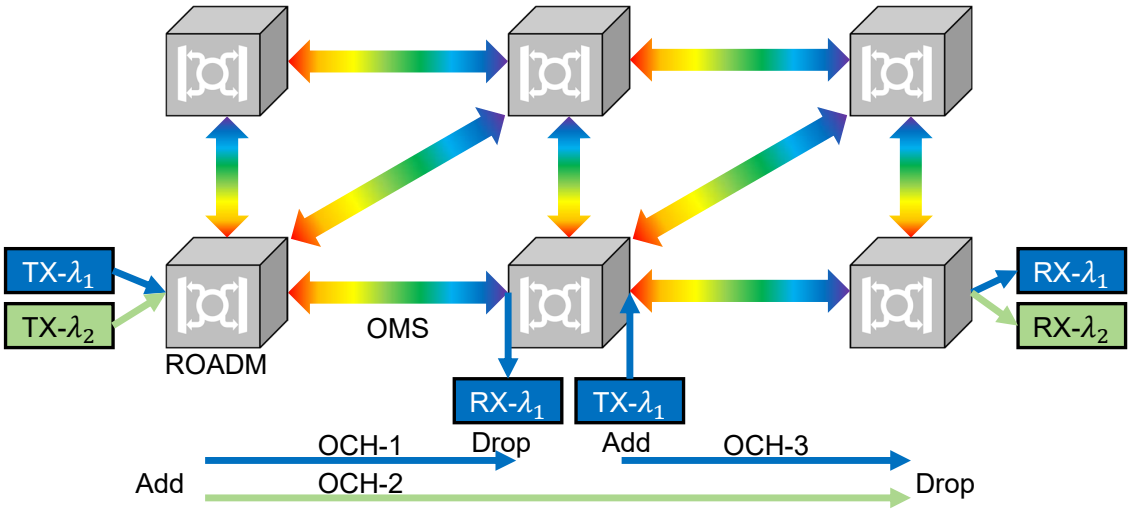


Figure 2.15: Sample optical transport network.

amplified multi-wavelength optical path across a fiber direction, while the ROADM layer selectively adds, drops, or bypasses individual OCh channels without disturbing the remaining wavelengths within the OMS. This separation of responsibilities allows operators to modify channel routing dynamically while maintaining continuous optical transmission.

A point-to-point WDM system can be viewed as consisting of two OMS endpoints connected by a single fiber pair with inline amplifiers. When ROADMs are inserted, each fiber direction becomes a "degree", and the WSS modules at a ROADM node connect the OMS of multiple degrees, enabling wavelength-level connectivity between any incoming and outgoing directions, as shown in Fig. 2.14. Modern ROADM architectures allow any wavelength to be added or dropped on any port (colorless), switched to any direction (directionless), and reused without wavelength contention (contentionless). As more degrees are added to ROADM nodes, the network topology transitions from linear or ring architectures into flexible mesh networks, as shown in Fig. 2.15.

In a ROADM-based mesh network, each node contains multiple OMS sections, each corresponding to a fiber direction. The WSS units interconnect these OMS sections at the OCh layer, forming an optical cross-connect that can dynamically route wavelengths across the network. The OMS amplifiers, power equalizers, Raman/EDFA modules, and optical supervisory channels ensure that all wavelengths belonging to an OMS are properly amplified and monitored, even as their routing paths change due to network demands or protection switching.

This architecture enables a number of key capabilities unavailable in point-to-point systems:

- **Survivability and protection:** Wavelengths can be rerouted around failures via alternate ROADM degrees, significantly improving resilience.
- **Efficient resource utilization:** Traffic can follow routing and wavelength assignment (RWA) paths, allowing for more scalable spectrum and fiber usage.

By leveraging OMS-based amplification and ROADM-based wavelength switching, operators can construct large-scale transparent optical meshes that support flexible path computation and automatic power balancing. Such ROADM-centric mesh networks greatly reduce operational complexity, enhance agility, and form the optical foundation for future high-capacity transport domains including metro-core, long-haul, and data-center interconnect networks.

Chapter 3

Quality of Transmission

Quality of transmission (QoT) in optical transmission systems is a critical metric that evaluates the performance and reliability of data transmission over optical fibers. Various parameters are included, such as signal-to-noise ratio (SNR), bit error rate (BER), and optical signal-to-noise ratio (OSNR). The QoT assessment is essential to ensure that transmitted data maintain integrity and meet required standards, especially in high-speed and long-haul transmission.

In this chapter, we will introduce the noises in optical transmission systems and the QoT based on them in Section 3.1, the Gaussian noise (GN) model and GN model-based optical power optimization in Section 3.3 and Section 3.4, and discuss the power control in optical networks in Section 3.5.

3.1 Noises and QoT Metrics in Optical Transmission Systems

3.1.1 ASE Noise and OSNR

In the Section 2.4.1, we introduced the ASE noise comes from EDFA. Owing to its stochastic nature and large number of independent noise contributions, ASE is commonly modeled as additive white Gaussian noise with a nearly flat spectral density across the amplifier gain bandwidth. The presence of ASE degrades the received signal quality, increases BER, and plays a central role in system design, as it directly determines the OSNR budget and transmission reach of long-haul coherent optical networks.

The power spectral density (PSD) of ASE noise is estimated in the form [49]:

$$N_{\mathrm{ASE}}(f) = n_{\mathrm{sp}} h f (G(f) - 1), \tag{3.1}$$

where n_{sp} is the spontaneous emission factor, h is the Plank's constant, f is the central frequency of the channel and $G(f)$ the gain of the amplifier.

Then, the ASE noise power is:

$$P_{\mathrm{ASE}}(f) = 2 N_{\mathrm{ASE}}(f) B_{\mathrm{ref}}, \tag{3.2}$$

where 2 is the factor of two polarizations and B_{ref} is the reference bandwidth.

Conventionally, the B_{ref} is considered as 12.5 GHz (0.1 nm in C-band) when computing the optical signal-to-noise ratio (OSNR):

$$OSNR(f) = \frac{P_{\mathrm{ch}}(f)}{P_{\mathrm{ASE}}(f)}, \tag{3.3}$$

where P_{ch} is the power of the optical channel in linear scale.

From Eq. (2.22), we can describe the NF of EDFA in linear scale as:

$$\mathrm{NF}(f) = 2 n_{\mathrm{sp}} \frac{G(f) - 1}{G(f)}. \tag{3.4}$$

3.1.2 Nonlinear Noise and GSNR

Nonlinear (NL) noise arises from the Kerr nonlinearity of optical fibers and represents a fundamental impairment in coherent WDM transmission systems operating at high launch powers. In dispersive fiber links, the interaction among wavelength-division multiplexed channels through self-phase modulation, cross-phase modulation, and four-wave mixing generates stochastic nonlinear distortions whose statistics, under typical long-haul conditions, are well approximated by additive Gaussian noise. This nonlinear interference accumulates along the link similarly to amplified spontaneous emission (ASE), but its magnitude strongly depends on the per-channel launch power, fiber parameters, channel spacing, and dispersion map. To quantify system performance in the presence of both ASE and NL noise, the generalized signal-to-noise ratio (GSNR) is widely adopted:

$$GSNR(f) = \frac{P_{\mathrm{ch}}(f)}{P_{\mathrm{ASE}}(f) + P_{\mathrm{NL}}(f)}. \tag{3.5}$$

GSNR consolidates linear and NL contributions into an equivalent end-to-end

SNR metric, defined as the ratio between the signal power at the receiver and the sum of ASE and NLI noise powers. As a result, GSNR provides a unified figure of merit that directly correlates with achievable information rates and bit-error performance, enabling accurate QoT estimation in optical transmission systems.

3.1.3 Transceiver Noise

Transceiver (TRX) noise encompasses all implementation-related impairments generated within the optical transmitter and receiver front-ends, including DAC/ADC quantization noise, driver and TIA noise, finite-resolution DSP operations, laser phase noise residuals, and modulator nonlinearities. Unlike ASE and fiber nonlinear interference, which accumulate along the optical link, TRX noise is local to the transceiver and often dominates performance in short-reach and metro coherent systems.

TRX noise is typically modeled as additive Gaussian noise whose power scales with the electrical signal level, leading to an effective implementation SNR, so-called SNR_{TRX}, that limits the maximum achievable SNR even in noise-free optical channels. Fig. 3.1 shows a back-to-back (b2b) characterization of a TRX.

The overall end-to-end SNR results from the combination of TRX noise, ASE noise, and nonlinear interference:

$$\frac{1}{SNR} = \frac{1}{GSNR} + \frac{1}{SNR_{\mathrm{TRX}}}, \tag{3.6}$$

or

$$NSR = GNSR + NSR_{\mathrm{TRX}}, \tag{3.7}$$

where NSR, $GNSR$, NSR_{TRX} are the noise-to-signal ratios.

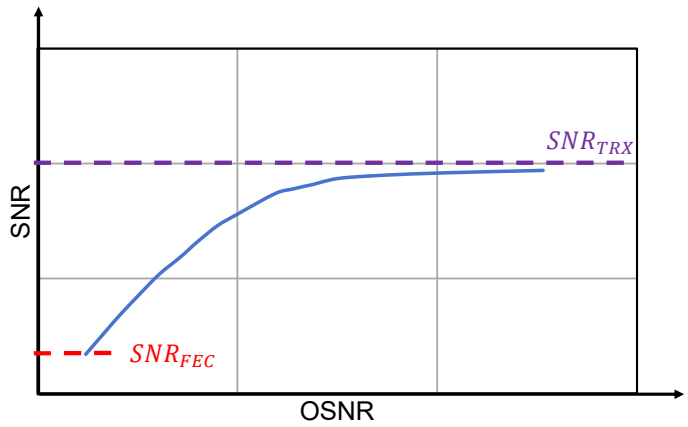


Figure 3.1: Sample b2b characterization curve of a TRX.

3.1.4 Other Impairments

Beyond ASE noise, nonlinear interference, and transceiver-related penalties, several additional impairments influence the performance of coherent optical transmission systems. Among these, filtering penalties arise from the cumulative effect of ROADM/WSS bandwidth constraints and the finite optical filtering stages in transceivers. Tight optical filters distort the signal spectrum, introduce inter-symbol interference, and increase linear crosstalk between adjacent WDM channels—especially in flex-grid networks where channels are packed with minimal guard bands. Repeated filtering in multi-hop mesh networks can further induce passband narrowing and phase ripple, which reduce the effective SNR and constrain achievable symbol rates and modulation formats.

Another important class of impairments includes polarization-dependent effects, namely polarization-dependent loss (PDL) and polarization-mode dispersion (PMD). PDL causes unequal attenuation of the two polarization components, leading to random variations of the received constellation and a reduction in the effective SNR, particularly when combined with nonlinearities and fading. PMD, caused by random birefringence variations in the fiber, introduces differential group delay between the two polarization states. Although coherent DSP can compensate a significant portion of PMD, rapidly varying or higher-order PMD leaves residual distortions and contributes to implementation penalties.

In practice, these secondary impairments are incorporated into system modeling through a combination of analytical approximations and implementation-specific penalty terms. Filtering penalties are often modeled by applying the concatenation of the actual ROADM/WSS transfer functions or by using equivalent low-pass filters that capture passband narrowing and phase ripple. PDL is typically represented by random Jones matrices with specified mean loss and statistics, allowing Monte-Carlo evaluation of its impact on SNR and GSNR. PMD is modeled as a stochastic sequence of birefringent sections or by drawing differential group delay (DGD) values from the Maxwellian distribution. Additional effects such as frequency drift, IQ imbalance, or residual dispersion mismatch are incorporated at the DSP level as linear impairments—such as phase rotation, frequency offset, skew, or amplitude imbalance—while crosstalk is modeled as additive interference with coherent or incoherent coupling depending on the source. Together, these modeling approaches enable the inclusion of "other impairments" within QoT estimation frameworks, either through direct simulation (SSFM or DSP-level models) or by aggregating them into effective implementation penalty terms used in GSNR-based performance prediction.

3.2 BER, SNR Margin and Shannon Capacity

In digital communication systems, three fundamental metrics govern the evaluation of performance: BER, SNR, and total capacity. The BER quantifies the probability that a transmitted bit is incorrectly detected at the receiver, thus representing the end-to-end quality experienced by the system. It depends on the modulation format, the underlying noise statistics, the receiver architecture, and physical-layer impairments. As a result, BER is often used as the primary figure of merit for system feasibility and as a direct constraint in QoT estimation in optical networks.

The BER performance is underlying by the SNR, defined as the ratio between the received signal power and the total noise power within the relevant detection bandwidth. The SNR provides an abstraction layer that separates the modulation-dependent detection process from the physical-layer noise contributions, such as ASE noise, nonlinear interference noise, and transceiver noise. For many modulation formats, analytical relationships exist between SNR and BER (for example, through Q-factor approximations for QAM constellations). These relationships enable SNR to serve as a convenient intermediate metric for performance prediction, system optimization, and network-level routing decisions.

Shannon capacity provides the theoretical upper bound on the information rate that can be transmitted error-free over a channel with a given SNR. For an additive white Gaussian noise (AWGN) channel, the Shannon capacity is defined as [50]:

$$C = R_s \log_2 (1 + \mathrm{SNR}) , \tag{3.8}$$

where R_s is the symbol rate.

Then the total throughput limit for a WDM system is:

$$C_{\mathrm{WDM}} = 2 \sum_{s=1}^{N_{\mathrm{ch}}} R_s(s) \log_2 (1 + \mathrm{SNR}(s)) , \tag{3.9}$$

where N_{ch} is the number of optical channels in the network.

Although practical modulation and coding schemes generally do not achieve this limit, the Shannon capacity acts as a benchmark for evaluating attainable spectral efficiency, guiding the design of FEC schemes, and assessing how close modern optical transmission systems operate to fundamental limits.

Here, we define the network-wide SNR margin as:

$$\mathrm{SNR}_{\mathrm{margin}} = \min_{s} (\mathrm{SNR}(s) - \mathrm{SNR}_{\mathrm{FEC}}(s)), \tag{3.10}$$

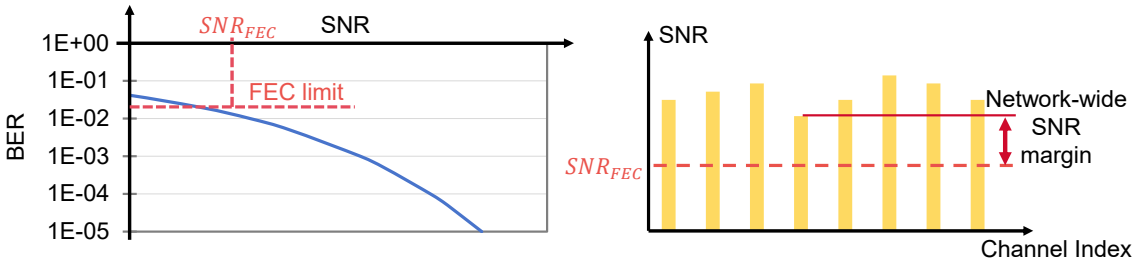


Figure 3.2: Network-wide SNR margin.

where $\mathrm{SNR}(s), s = 1, ..., N_{\mathrm{svc}}$ is the SNR of services s, and $\mathrm{SNR}_{\mathrm{FEC}}(s)$ is the SNR at the FEC limit of the transponder for service s. The SNR margin represents the performance of BER [9, 51]. If $\mathrm{SNR}(s) < \mathrm{SNR}_{\mathrm{FEC}}(s)$, then uncorrected blocks will appear. Therefore, the SNR margin should always be positive during any optical network operation. The higher SNR margin, the lower pre-FEC BER, which can then be leveraged to carry additional traffic / increase the network capacity.

3.3 Gaussian Noise (GN) Model

The Gaussian noise (GN) [52] model provides an analytical foundation for estimating the nonlinear interference (NLI) generated in WDM systems. Its principle is based on the observation that in dispersion-uncompensated transmission links, the large accumulated chromatic dispersion causes the temporal and spectral components of the WDM channels to decorrelate rapidly. This decorrelation leads the Kerr-induced nonlinear mixing terms to exhibit broadband Gaussian statistical behavior. By adopting a perturbative treatment of the NLSE, the GN model describes nonlinear distortions as an additive, signal-independent complex Gaussian noise term whose power spectral density (PSD) can be written in closed form. This approximation is valid under weak nonlinearity, large accumulated dispersion, and the assumption that the transmitted waveform behaves as a wide-sense stationary Gaussian process. Because dispersion destroys phase coherence between fiber spans, the NLI generated in each span accumulates incoherently, allowing NLI contributions to be added at the power level across spans and optical sections.

The enhanced Gaussian noise (EGN) model extends the classical GN model by relaxing the Gaussianity assumption. While the GN model treats the channel as a Gaussian process and therefore fails to accurately represent the nonlinear behavior of finite-cardinality modulation formats such as QPSK, 16QAM, the EGN model incorporates the exact higher-order moments of the constellation, including kurtosis, to capture modulation-format-dependent nonlinear distortions. In addition,

the EGN formulation accounts for partial nonlinear coherence effects that become relevant when the accumulated dispersion per span is not sufficiently large or when Raman amplification, tight channel spacing, or metro-reach spans alter the spectral distribution of nonlinear interactions. As a result, the EGN model improves accuracy to within sub-dB deviation from split-step Fourier method (SSFM) simulations while maintaining a significantly lower computational cost than full numerical propagation.

Hereafter, we only use the item "GN" to represent the methods of GN and EGN.

The applicability of the GN model extends naturally into the hierarchical impairment evaluation framework used in optical transport networks, where QoT must be assessed from the OMS level down to the end-to-end optical channel (OCh). At the OMS level, the nonlinear noise generated by each WDM channel is computed using the GN model based on the span parameters, including fiber dispersion, nonlinear coefficient, span loss, gain profile of the amplifiers, channel spacing, and launched optical power.

Based on the GN model [14], the GSNR of a channel in an OMS is [53]:

$$GSNR_{\mathrm{OMS}} = \frac{P_{\mathrm{ch}}}{P_{\mathrm{ASE}} + P_{\mathrm{NL}}} = \frac{P_{\mathrm{ch}}}{P_{\mathrm{ASE}} + \eta P_{\mathrm{ch}}^3} \tag{3.11}$$

where P_{ch} is the channel power, P_{ASE} is the ASE noise power, P_{NL} is the NL noise power, which is proportional to P_{ch}^3 with a ratio η.

Once the OMS-level impairments are computed, QoT estimation progresses to the OCh level by aggregating impairments along the entire route of the optical channel. Under the incoherent accumulation assumption of the GN model, the end-to-end NL power for channel is obtained by summing the NL contributions from all OMS sections along its light path (LP). A typical OCh may traverse multiple OMS segments between the transmitter and receiver. For a service s carried by wavelength λ through LP $\mathrm{OMS}_1 - \mathrm{OMS}_2 - \ldots - \mathrm{OMS}_{\mathrm{N}}$, the end-to-end (e2e) GSNR is [15]:

$$GSNR_{\mathrm{e2e}}(s) = GNSR_{\mathrm{e2e}}^{-1}(s) = \left(\sum_{n=1}^{N} GNSR_{\mathrm{OMS}_n} \right)^{-1}, \tag{3.12}$$

where GNSR is the generalized noise-to-signal ratio, that is the inverse of GSNR.

Adding WSS distortion and transponder noises [9, 54], the e2e SNR can be written as:

$$SNR = NSR^{-1} = (f_{\mathrm{WSS,TRX}}(GNSR) + NSR_{\mathrm{TRX}})^{-1}, \tag{3.13}$$

where $f_{\mathrm{WSS,TRX}}$ represents the filtering penalty from WSSs and noise scaling in TRX, NSR_{TRX} represents the noise introduced by TRX, which can be calibrated in back-to-back (B2B) performance measurements.

SNRs estimated or measured throughout this paper are inherently end-to-end, capturing the cumulative impact of all the optical network elements on signal quality.

This hierarchical impairment estimation, starting from the physical characterization of each OMS and culminating in the end-to-end OCh-level QoT, forms the foundation of physical-layer-aware routing, wavelength assignment, and network optimization in modern coherent ROADM networks. The GN model offers the computational efficiency necessary for large-scale network evaluation, while the EGN model provides enhanced accuracy in scenarios where modulation-format dependence or partial coherence effects become significant. SSFM simulations continue to serve as the reference for validating analytical models but remain impractical for real-time network planning due to their high computational complexity.

3.4 Launch Power Optimization

Optimizing the optical launch power is a fundamental step in the design and operation of coherent optical transmission systems, as it directly determines the balance between linear noise accumulation and nonlinear distortion. At low launch powers, the signal is dominated by amplifier noise—primarily ASE—resulting in a poor OSNR. As the launch power increases, the OSNR initially improves; however, once the fibre's nonlinear regime is approached, Kerr-induced impairments such as SPM, XPM, and FWM begin to dominate, causing NLI that degrades system performance. The combined effect leads to a characteristic "nonlinear Shannon limit" where the achievable information rate peaks at an optimal input power.

Several methods have been developed to determine or approximate this optimal launch power. The most classical approach relies on the GN model or its enhanced versions (EGN), where nonlinear interference is treated as an additive Gaussian noise term. By analytically expressing NLI as a function of launch power, fiber dispersion, span length, and WDM configuration, the GN family of models enables closed-form estimation of the optimal per-channel power that maximizes the GSNR or the achievable information rate.

Based on Eq. (3.11), GSNR varies with channel power, as shown in Fig. 3.3. By calculating the derivative of Eq. (3.11), GSNR maximization is achieved by balancing the ASE-to-NL noise ratio to 3dB [14], therefore, the optimal launch

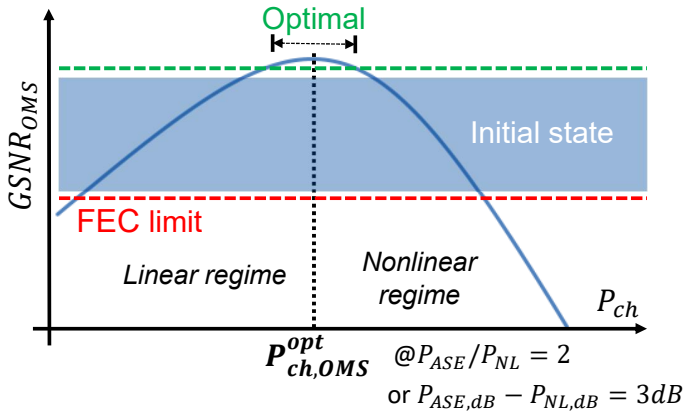


Figure 3.3: GSNR vs. launch power in an OMS.

power of the channel in this OMS is:

$$P^{\text{opt}}_{\text{ch,OMS}} = \sqrt[3]{\frac{P_{\text{ASE}}}{2\eta}}. \tag{3.14}$$

These models are widely used for offline network design, link engineering, and rapid QoT estimation due to their low computational complexity.

More accurate but computationally heavier approaches depend on numerical simulations, such as the SSFM. SSFM directly integrates the nonlinear Schrödinger equation and thus accounts for effects not captured by analytical models, including modulation-format dependence, signal statistics, and complex spectral shaping. Launch-power optimization using SSFM is typically applied in system research or validation studies where highest accuracy is needed. Machine-learning-based optimization has also emerged, where neural-network regressors or Bayesian optimizers learn the complex mapping between launch power and QoT metrics from training data, enabling adaptive optimization in dynamically changing networks.

In practical network operation, launch-power optimization can also be performed in real time. Modern coherent transceivers estimate metrics such as GSNR, nonlinear penalties, and equalizer error, which can serve as feedback for adaptive power control. Network-wide control planes may adjust power jointly with gain settings of EDFAs and Raman pumps to meet end-to-end QoT thresholds while reducing nonlinear penalties. In elastic and multi-band systems, optimization may be further combined with bandwidth allocation, modulation-format selection, and power pre-emphasis to mitigate inter-channel nonlinearities and filter-shaping penalties.

Overall, launch-power optimization methods form a spectrum ranging from analytical GN-based calculations to high-fidelity simulations and data-driven or closed-loop adaptive control. These approaches aim to operate each channel near its non-

linear optimum across all network conditions, ensuring maximized reach, spectral efficiency, and energy efficiency in modern coherent optical communication systems.

3.5 Power Control of Optical Networks

Power control in WDM optical networks is fundamentally driven by the interaction between amplifier physics, transmission impairments, and network dynamics. In ROADM-based optical networks, dynamic operations such as channel add/drop and rerouting introduce significant power fluctuations, as shown in Fig. 3.4, making power control a complex network-wide challenge.

3.5.1 Spectral Power Fluctuations in WDM Networks

Spectral power fluctuations in WDM systems arise from wavelength-dependent gain and nonlinear interactions among channels. Two dominant physical mechanisms contributing to these fluctuations are gain spectral hole burning (SHB) in optical amplifiers and stimulated Raman scattering (SRS) in optical fibers [21].

SHB occurs due to inhomogeneous saturation of the EDFA gain medium. As channels at specific wavelengths deplete the local population inversion, the gain spectrum becomes dependent on the instantaneous wavelength configuration. This results in wavelength-dependent gain offsets when channels are added or removed, even when the amplifier maintains a constant average gain. Experimental studies have shown that such gain offsets can reach several tenths of a dB and accumulate over multiple cascaded amplifiers [55].

In addition to SHB, SRS induces power transfer from shorter wavelengths to

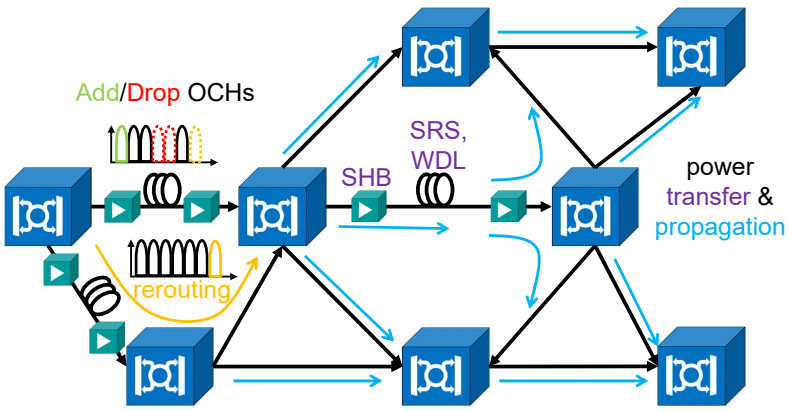


Figure 3.4: Power fluctuations due to channel add/drop and rerouting. SHB: spectral hole burning; SRS: stimulated Raman scattering; WDL: wavelength-dependent loss.

longer wavelengths during propagation in optical fibers. This effect creates a spectral tilt, which depends on the channel loading and transmission distance. As discussed in [23], the interplay between amplifier gain ripple and SRS-induced tilt leads to non-uniform power evolution across channels, further exacerbating spectral imbalances.

These effects collectively introduce a strong dependence of channel power on the global spectral configuration, making power control inherently a coupled and network-wide problem rather than an independent per-channel adjustment.

3.5.2 Short-term and Long-term Dynamics

The dynamics of power fluctuations in optical networks can be broadly categorized into short-term and long-term effects, depending on the time scale of the underlying physical processes and control mechanisms.

Short-term dynamics are primarily governed by the transient response of optical amplifiers. When channels are added or dropped, EDFAs experience rapid gain changes due to variations in total input power and population inversion. These transients typically occur on microsecond to millisecond time scales and can manifest as overshoots or undershoots in channel power. Various control techniques, such as automatic gain control (AGC) and feedforward mechanisms, have been developed to suppress these fast transients. For example, fast electrical feedforward control combined with variable optical attenuators (VOAs) can effectively reduce transient excursions without inducing oscillatory responses in cascaded amplifier chains [55].

In contrast, long-term dynamics correspond to steady-state power deviations arise from wavelength-dependent gain characteristics and channel power coupling effects. As shown in [23], the steady-state power of a channel depends on the spectral distribution of all other channels due to the total power control mechanism of EDFAs. Channel addition or removal events can therefore induce power excursions of several dB, which propagate across multiple spans and accumulate along the transmission path. Importantly, these long-term dynamics are not solely determined by the number of channels but also by their spectral positions and relative gain differences. This highlights the necessity of considering spectral configuration in both system design and real-time control.

3.5.3 Power Control Strategies

To mitigate spectral power fluctuations and ensure stable operation, various power control strategies have been proposed at both the device and network levels.

At the device level, EDFAs employ control schemes such as constant gain control or constant output power control. While these approaches stabilize the total power, they inherently introduce coupling among channels, as the control loop operates on aggregate power measurements rather than individual wavelengths [22]. GFFs are also used to reduce wavelength-dependent gain variations. However, residual gain ripple remains and can vary with operating conditions, limiting the effectiveness of purely static compensation techniques.

Given the limitations of local control, network-aware approaches have been proposed to coordinate power control across multiple nodes. These methods consider RWA, and power control jointly to minimize power fluctuations. For instance, gain-aware wavelength assignment strategies can reduce the impact of gain offsets by selecting channel configurations that minimize gain discrepancies [55]. Similarly, earlier works [56, 57] have demonstrated that incorporating physical-layer constraints into network control algorithms significantly improves system stability.

Recent research has also focused on the stability of distributed power control systems. Using control-theoretic and game-theoretic formulations, it has been shown that improper coordination among control loops can lead to oscillations or instability in mesh networks. The work in [58] highlights the importance of convergence properties and equilibrium analysis in distributed optical power control. These approaches provide a theoretical foundation for designing stable and scalable control mechanisms in future dynamic optical networks.

Chapter 4

Digital Twins of Optical Networks and Experimental Demonstrations

This chapter is based on the publications during the PhD [24, 26–28], Section 4.1 introduces the digital twins, Section 4.3 presents the power equalization problem in optical networks and Section 4.4 gives the proposed solutions based on our DT. The experimental setups used in this work are presented in Section 4.6 and the results shown in Section 4.7.

4.1 Digital Twins and Closed-loop Control

Digital twins (DTs), which are software replica of real physical systems, which can directly interact with the underlying physical system, have been proposed and used in optical networks for automation and management [59–62]. By monitoring its twin physical system, a DT can analyze the behavior of the underlying physical system using physical or machine learning models (or a combination of the two [8]). Then, the DT can be used as a sandbox to predict the impact of changes to the system through emulation within the DT, thereby improving decisions before implementing any operation in the physical world. A typical workflow using a DT is shown in Fig. 4.1.

4.2 AI-Light: a Platform for Optical Networks Automation

Software-defined networking (SDN) [63, 64] enables a global network view for monitoring and actuation allowing the use of autonomous applications that enhance

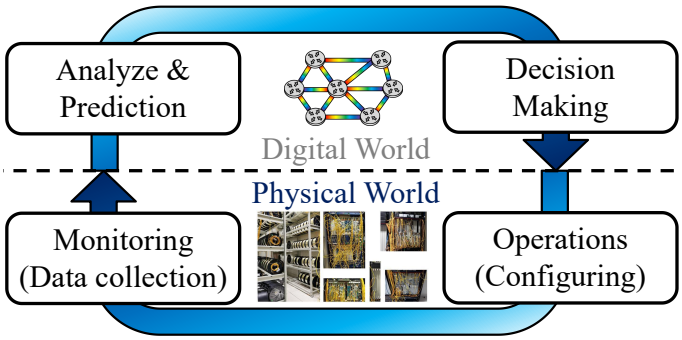


Figure 4.1: DT-enabled closed-loop control.

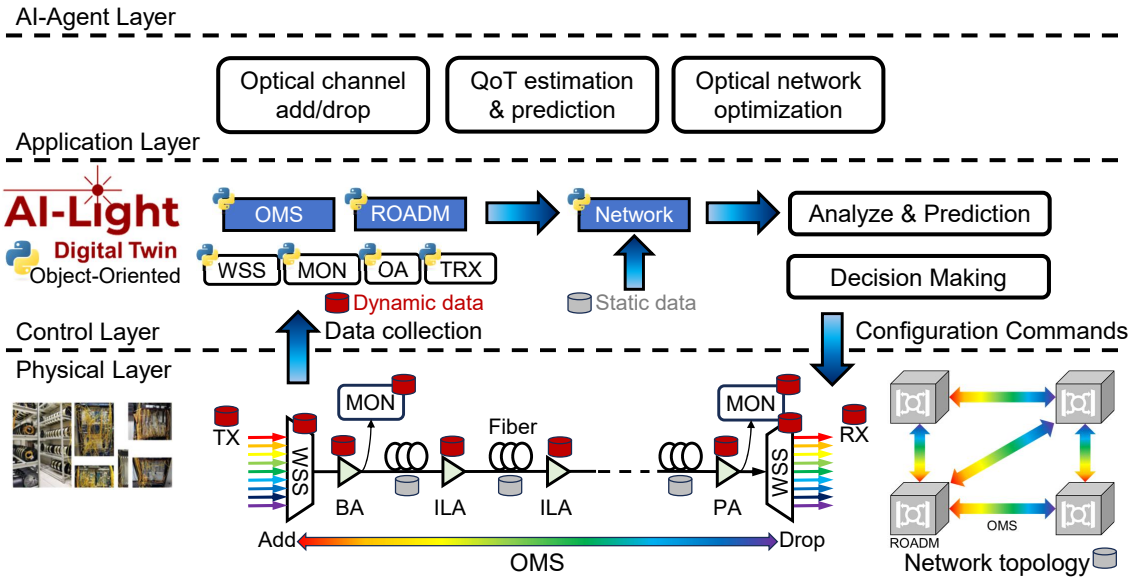


Figure 4.2: AI-Light architecture.

network performance in terms of capacity and quality.

Network intelligence is logically centralized in software-based SDN controllers that maintain a global view of the network, which appears to applications and policy engines as a single, logical switch. Network control is directly programmable in SDN since it is decoupled from forwarding functions. SDN lets network managers configure, manage, secure, and optimize network resources very quickly via dynamic, automated SDN programs.

AI-Light is a proprietary Python-based SDN framework developed in our lab. It demonstrates a complete automation cycle as shown in the Fig. 4.2: the telemetry is leveraged to probe the current state of the network and the retrieved information is combined with the static information stored in a database to create a reliable DT (a network state example is presented in the appendix A); an SNR-based optimization algorithm computes the optimum configuration over the DT by leveraging a quality-of-transmission estimator (QoT-E); the configuration of the WSSs is used to

achieve the optimum power per channel and the amplifiers are configured in gain-lock mode to simplify the management of fiber cuts and the deployment of new services. Finally, the optimized configuration is pushed to the physical network equipment.

The physical layer (or infrastructure layer) connects the network elements (NEs) in the lab so that we can perform the measurements, the monitoring, the configurations from top to the bottom.

The control layer provides classes (Network, OMS, ROADM, etc.) that can be used to create the required objects as interfaces to control the actual network and carry out different tasks.

The application layer of AI-Light allows network configuration, OMS configuration, ROADM configuration, automatic power configuration, DTs generation, lightpath establishment, power optimization, etc. by the end-user or operator.

Last but not least, the AI-Agent layer leverages the LLMs will be introduced in Chapter 5.

4.3 Problem of Power Equalization in Optical Networks

Methods to optimize optical networks through per-channel launch power setting (also known as power equalization) based on physical models [14–18] or machine learning models [19, 20] have been proposed and widely applied. SNR, as a significant criterion to assess the QoT of a communication system, can be improved thanks to power equalization. However, when networks are operating for a long time, e.g., in "set and forget" mode or after unforeseen events that change the underlying physical layer (e.g., span loss increase after repairing a fiber cut), power settings hence SNR may become suboptimal [13]. For this reason, it is important to periodically re-optimize the network.

Power equalization practically consists of adjusting the power of one or several channels of one or more OMSs through WSS per-channel attenuation change. Changing several channels' power on a single WSS is possible; however, simultaneously changing the settings of WSS on several OMSs, or even on the same OMS, is virtually impossible, such that WSS attenuation profile change can only be considered as sequential rather than parallel operations, and the impact of the change of one WSS attenuation profile, on OMSs downstream of said WSS, always needs to be considered when equalizing powers.

The spectral power fluctuation due to different effects has been discussed in Section 3.5.1.

Specifically, when equalizing the power of a multi-OMS service, it is possible that the SNR of this service actually decreases during the equalization process, before increasing again and reaching the desired value. This may be caused by nonlinear inter-channel noise or power transfers phenomena across the channels when power changes on one OMS propagates to downstream OMSs [21, 27, 28].

Commercial optical amplifiers can be set in either gain-lock or power-lock mode. While the power-lock could prevent undesired power propagation on downstream links, it also hinders fast rerouting after a failure; indeed, if services are rerouted on an OMS set to power-lock mode, the total OMS launch power will increase leading to a per-service power decrease (after the booster of the OMS) and hence to suboptimal power allocation not only for the rerouted services, but also for the existing services. Operating amplifiers in gain-lock mode avoids this problem.

Hence, network-wide re-optimization and autonomous optical networking require the ability to search for a sequence of power adjustments whereby existing channels' SNRs do not degrade [65–67]. We propose to use an optical network DT as a sandbox to search for such a safe and fast sequence of operations. The DT implements real-time performance monitoring and SNR estimation/prediction. An accurate SNR estimator requires accurate knowledge of the physical parameters. In real networks, the values of those parameters may be unknown, incorrect, or outdated, for instance when powers are changed during the re-optimization process. Online monitoring and updating of the physical parameters to close the monitoring–decide–act control loop is needed for accurate SNR prediction [65, 66].

4.4 Methodology: DT-enabled Autonomous Power Equalization

The methodology we propose here is generic and independent of the specific power equalization strategy used in the network, see Fig. 4.3: We build the DT and periodically update it based on physical parameters monitored in the network.

$$f(\theta) : Net \rightarrow \widehat{Net}, \tag{4.1}$$

where θ represents the physical parameters from the real world, including power spectra, gain/tilt of OAs, fiber parameters in the OMSs, topology and services information of network, etc. Later, we use hat to represent the processes in the

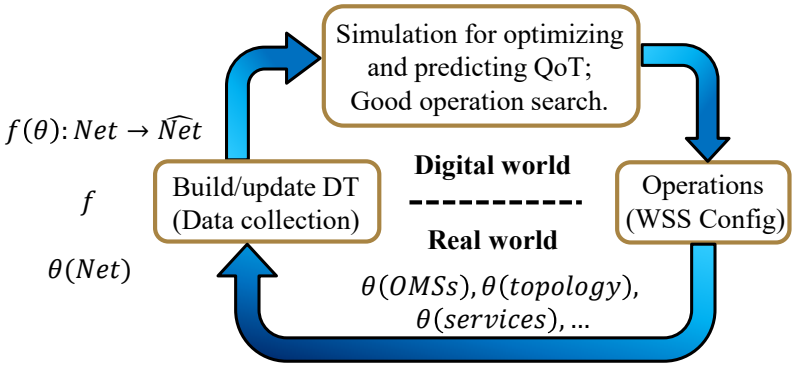


Figure 4.3: DT enabled power equalization.

digital world.

Then, we find the order of the operations along with the power variation step-size, and assess the impact of those changes on the QoT of all services using a QoT estimator, before implementing the operations in the network to prevent potential QoT degradations.

4.4.1 Power Propagation in a Mesh Network Digital Twin

The spectral power fluctuation due to different effects is discussed in Section 3.5.1 and it is necessary to simulate the power propagation before operating on the real network, whenever we want to adjust some power settings. To efficiently emulate power propagation in a mesh optical network, we propose a heuristic approach that decomposes the network topology into a set of ordered propagation chains (Algorithm 1). This method enables sequential and localized power updates (Algorithm 2) while ensuring that the impact of a power variation at a source OMS is propagated throughout the entire network (Algorithm 3).

Chain-Based Network Decomposition

The optical network is modeled as a graph $G = (V, E)$, where each node represents an OMS, and each edge represents the connection between two OMSs.

Starting from a source OMS src, a depth-first search (DFS) is used to enumerate all non-cyclic propagation chains, see Algorithm 1:

$$\mathcal{C}_k = (src, v_1, v_2, \ldots, v_n) \tag{4.2}$$

such that no edge is revisited during traversal. The resulting set of chains $\{\mathcal{C}_k\}$ defines all possible propagation paths originating from the source.

Algorithm 1: DFS-based chain search from source OMS

Input: Network graph $G = (V, E)$; source OMS src.
Output: Set of propagation chains $\mathcal{P}$.

Initialize stack with $(\text{None}, src, [\,])$
Initialize visited edge set $\mathcal{S} \leftarrow \emptyset$
Initialize chain set $\mathcal{P} \leftarrow \emptyset$
while *stack not empty* **do**
    Pop $(u, v, path)$ from stack
    Append v to $path$
    Mark edge (u, v) as visited in $\mathcal{S}$
    Find neighbors $w \in \mathcal{N}(v)$ such that $(v, w) \notin \mathcal{S}$ and $w \neq src$
    if *no such neighbor exists* **then**
        Append $path$ to $\mathcal{P}$
    else
        for *each valid neighbor* w **do**
            Push $(v, w, path)$ to stack
return $\mathcal{P}$

Algorithm 2: Power propagation along a chain

Input: DT physical model $\mathcal{M}$; chain $\mathcal{C}_k$.
for *each OMS* $i \in \mathcal{C}_k$ **do**
    Retrieve initial per-channel output power $P_{i,\lambda}^{(0)}$ at pre-amplifier;
    if $i \neq src$ **then**
        Update transmitted channels at booster:

$$P_{i,\lambda} \leftarrow P_{i,\lambda} + \Delta P_\lambda$$

    Compute updated output power at pre-amplifier $P_{i,\lambda}^{(1)}$ in the $\mathcal{M}$;
    if *next OMS* j *exists for channel* λ **then**
        Identify transmitted channel set $\Lambda_{i \to j}$
        Compute output power variation at pre-amplifier:

$$\Delta P_\lambda = P_{i,\lambda}^{(1)} - P_{i,\lambda}^{(0)}, \quad \forall \lambda \in \Lambda_{i \to j}$$

Sequential Power Propagation Along a Chain

For each chain $\mathcal{C}_k$, power is propagated sequentially from upstream to downstream OMSs. At each OMS, the DT physical model is updated and the variation of per-channel power is computed and forwarded to the next OMS along the chain, as shown in Algorithm 2.

Algorithm 3: Global chain-based power propagation

Input: DT with network graph G and physical model $\mathcal{M}$; power perturbation source OMS *src*.

1 Power adjustment at *src*;
2 Search all propagation chains $\mathcal{P} \leftarrow$ `SearchChains`(G, src)
3 **for** *each chain* $\mathcal{C}_k \in \mathcal{P}$ **do**
4 Convert chain representation (e.g., OMS labels to indices)
5 Apply power propagation: `Propagate`$(\mathcal{M}, \mathcal{C}_k)$
6 **return** *Updated DT.*

Global Power Propagation from Source OMS

To emulate the effect of a power adjustment at the source OMS *src*, the propagation procedure is applied to all chains originating from *src*. This corresponds to exhaustively traversing all propagation light paths and executing the propagation function along each of them, as shown in Algorithm 3.

This exhaustive traversal ensures that all OMSs reachable from the source are updated, allowing the DT to emulate the network-wide impact of a local power variation. Since multiple chains may share common OMS segments, certain OMSs can be updated multiple times within a single propagation cycle, reflecting the multi-path nature of power interactions in mesh optical networks.

Overall, this method can be interpreted as a single-pass approximation of a distributed fixed-point iteration, providing a practical trade-off between computational efficiency and physical accuracy for real-time DT applications.

4.4.2 Multi-Step Lookahead Prediction with Fixed Step-Size

Multi-step lookahead in chess is the ability to anticipate and evaluate potential move sequences several turns ahead, considering both the player's and the opponent's responses. This strategic foresight enables players to plan, exploit opportunities, and avoid risks effectively.

In the power equalization problem, the multi-step lookahead method can avoid local optimum step-searching thereby reducing the total time consumption.

Algorithm 4 presents the workflow of multi-step lookahead prediction and configuration of an optical network with fixed step-size.

First, data collection is needed for building/updating DT (line 2). However, it is not necessary to update the DT after each step. Instead, we update our DT every K_{update}-steps (line 4).

Second, assuming there are N_{OMS} non-optimized $\widehat{OMS}$, the optimal launch

Algorithm 4: Multi-step lookahead prediction for power equalization with a fixed step-size

Input: Initial DT model; optimization conditions (Condition-1/2/3); parameters K_{update}, K_m, N_{OMS}, and $N_{\text{CH(n)}}$.

1 **while** *NOT optimized for all* $\widehat{OMS}$ *(Condition-1/2/3)* **do**
2 Update DT.
3 Find the optimal launch power per $\widehat{OMS}$ by Eq. (3.14).
4 **for** $k = 1$ *to* K_{update} **do**
5 **for** $m = 1$ *to* K_m **do**
6 **for** $n = 1$ *to* N_{OMS} **do**
7 **for** $\lambda = 1$ *to* $N_{CH(n)}$ **do**
8 Adjust booster $\widehat{P_n}(\lambda)$ by Eq. (4.3).
9 Power propagation (Alg. 3) in DT.
10 $\widehat{SNR}$ prediction in DT.
11 Find the K_m-step order of operations which yields the highest $\widehat{SNR}_{\text{margin}}$ by Eq. (4.6).
12 Configure the WSS.
13 **return** *Optimized configuration of* $\widehat{OMS}$, $\widehat{P_n}(\lambda)$, *and* $\widehat{SNR}_{margin}$.

power profiles of these $\widehat{OMS}$s are found by DT based on Eq. (3.14) $\widehat{P^{\text{opt}}_{\text{ch,n}}}(\lambda)$ (line 3)
.

Third, for K_m-step lookahead (lines 5-11), the DT emulates the power adjustment impact on different $\widehat{OMS}$s (line 6). The fixed step adjustment (line 8) of launch power $\widehat{P_{\text{ch,n}}}(\lambda)$ of each channel of n^{th} $\widehat{OMS}$ given by:

$$\Delta\widehat{P_{\text{n}}}(\lambda) = \begin{cases} \widehat{P^{\text{opt}}_{\text{ch,n}}}(\lambda) - \widehat{P_{\text{ch,n}}}(\lambda) & \text{if } |\text{Err}(\widehat{P_{\text{ch,n}}}(\lambda))| < \delta, \\ \delta \cdot \text{sign}(\text{Err}(\widehat{P_{\text{ch,n}}}(\lambda))) & \text{otherwise.} \end{cases} \tag{4.3}$$

where δ is a fixed, predefined power adjustment step-size, $\text{Err}(\widehat{P_{\text{ch,n}}}(\lambda))$ is the power error.

We defined the power error between the current channel power $P_{\text{ch,n}}(\lambda)$ and the target optimized power $P^{\text{opt}}_{\text{ch,n}}(\lambda)$ of the n^{th} OMS:

$$\text{Err}(P_{\text{ch,n}}(\lambda)) = P^{\text{opt}}_{\text{ch,n}}(\lambda) - P_{\text{ch,n}}(\lambda). \tag{4.4}$$

The average absolute power error $Avg(|\text{Err}(P_{\text{ch,n}})|)$ shows how far the current state is from the optimized (target) state.

Then, the DT predicts the $\widehat{SNR}$ (line 10) after the power propagation (line 9)

based on Alg. 3.

For K_m-step lookahead within a N_{OMS}-OMS network, there are up to $N_{\text{OMS}}^{K_m}$ possible sequences (this is an upper bound because if an operations decrease the margin at some step, then the algorithm will not pursue the branch):

$$seq\{k_i\} = (\widehat{OMS}_{k_1}, \ldots, \widehat{OMS}_{k_i}, \ldots, \widehat{OMS}_{k_{Km}}), \tag{4.5}$$

where $i \in \{1, 2, ..., K_m\}$, and $k_i \in \{1, 2, ..., N_{\text{OMS}}\}$.

For each sequence, line 10 yields $\Delta\widehat{SNR}_{seq\{k_i\}}$ after power propagation. Therefore, the DT picks up the order of operations (on which $\widehat{OMS}$) yields the highest $\widehat{SNR}_{\text{margin}}$ (line 11):

$$\begin{cases} \max\limits_{seq\{k_i\}} \Delta\widehat{SNR}_{seq\{k_i\}}, \\ \Delta\widehat{SNR}_{k_i} > \epsilon_{\text{SNR}}, \end{cases} \tag{4.6}$$

where ϵ_{SNR} is the tolerance of $\widehat{SNR}$ drop during the multi-step operations. To avoid oscillations/local optima, the multi-step lookahead algorithm allows small $\widehat{SNR}$ drops at intermediate steps k_i in case this yields a higher $\widehat{SNR}$ after all K_m steps.

We use the power error metric as the most important convergence criterion for the optimization algorithm. The SNR margin quantifies the performance of the worst service in the network, and should remain (as much) positive (as possible). The capacity C quantifies the global performance of all services in the network.

Then, the DT will configure the WSS (line 12) following the order from line 11, until all $\widehat{OMS}$s are optimized (line 1).

Convergence

The convergence of the algorithm, or the stop-while condition, could be defined by using the metrics in Eq. (4.4-3.9):

Condition-1: average power error tolerance ε_{Perr}.

$$\text{Avg}(|\text{Err}(\widehat{P_{\text{ch,n}}})|) \leq \varepsilon_{Perr}. \tag{4.7}$$

Condition-2: max SNR margin error from ideal value $\varepsilon_{\text{margin}}$.

$$|\widehat{SNR}_{\text{margin}} - \max_{th}(\widehat{SNR}_{\text{margin}})| \leq \varepsilon_{\text{margin}}, \tag{4.8}$$

where $\max_{th}(\widehat{SNR}_{\text{margin}})$ is the max theoretical SNR margin achievable in DT.

Algorithm 5: Multi-step lookahead prediction for power equalization with dynamic step-size

Input: Initial DT model; optimization conditions (Condition-1/2/3); parameters K_{update}, K_m, N_{OMS}, $N_{\text{CH(n)}}$, and step-size range $\delta_{max} \rightarrow \delta_{min}$.

while *NOT optimized for all* $\widehat{OMS}$ *(Condition-1/2/3)* **do**
  Update DT.
  Find the optimal launch power per $\widehat{OMS}$ by Eq. (3.14).
  for $k = 1$ *to* K_{update} **do**
    for $m = 1$ *to* K_m **do**
      for $n = 1$ *to* N_{OMS} **do**
        for $\delta = \delta_{max}$ *down to* δ_{min} **do**
          for $\lambda = 1$ *to* $N_{CH(n)}$ **do**
            Adjust booster $\widehat{P_n}(\lambda)$ by Eq. (4.3) and Eq. (4.10).
          Power propagation (Alg. 3) in DT.
          $\widehat{SNR}$ prediction in DT.
    Find the K_m-step order of operations which yields the highest $\widehat{SNR}_{\text{margin}}$ or maximizes power error reduction Eq. (4.11).
  Configure the WSS.
return *Optimized configuration of* $\widehat{OMS}$, $\widehat{P_n}(\lambda)$, δ, *and* $\widehat{SNR}_{margin}$.

Condition-3: capacity (overall SNR) error from ideal value ε_C.

$$|\widehat{C} - \max_{th}(\widehat{C})| \leq \varepsilon_C, \tag{4.9}$$

where $max_{th}(\widehat{C})$ is the max theoretical capacity achievable in DT.

These conditions can be used alone or in combination.

4.4.3 Dynamic Step-Size for Power Adjustment

As we presented in [26], when the optical network was close to the optimized state, the measured SNR_{margin} was higher and increased slowly or even oscillated. To speed up convergence, a larger step-size could be considered and SNR_{margin} might no longer be the best metric when operations are selected in the DT. Then, we propose Algorithm 5.

There are two differences compared with Alg. 4.

1) Given a max step-size δ_{max}, a min step-size δ_{min} and interval $\Delta\delta$, the step-size

is dynamically adjusted (line 7 in Algorithm 5) within the range:

$$\delta = \delta_{min}, \delta_{min} + \Delta\delta, ..., \delta_{max} - \Delta\delta, \delta_{max}. \tag{4.10}$$

The loop starting line 7 predicts the $\widehat{SNR}$ after changing the launch power by a step-size in the above range. A larger step-size yields more power fluctuation and has higher probability to decrease the $\widehat{SNR}_{\text{margin}}$. Hence, the algorithm dynamically adjusts the step-size, starting from a large value, to find the max step-size δ ensuring safety yet minimizing the number of operations. For each step in line 9, we apply the above δ into Eq. (4.3).

2) The difference between line 11 in Alg. 4 and line 12 in Alg. 5 lies in the action search order; once the $\widehat{SNR}_{\text{margin}}$ is higher enough, DT will try to search the step orders which improves the $\widehat{SNR}_{\text{margin}}$ and reduces more power error $\Delta Avg(|\text{Err}(\widehat{P_{\text{ch,n}}})|)_{seq\{k_i\}}$ between current state profile and target state, instead of focusing on $\widehat{SNR}_{\text{margin}}$ only.

$$\begin{cases} \Delta\widehat{SNR}_{seq\{k_i\}} > 0, \\ \max\limits_{seq\{k_i\}} \Delta Avg(|\text{Err}(\widehat{P_{\text{ch,n}}})|)_{seq\{k_i\}}. \end{cases} \tag{4.11}$$

4.4.4 Parallel Configuration

Without prediction of $\widehat{SNR}$, the parallel configuration is risky for the optical network since the commands cannot be guaranteed to arrive at all network elements synchronously; intermediate states may interrupt services.

Instead, based on the multi-step lookahead, we can send the commands for K_m operations at the same time even if the WSS adjustments will not be strictly simultaneous, since prior emulation within the DT ensures no $\widehat{SNR}$ degradation for any service, whatever the order in which the commands are actually received by the physical layer from the DT, and executed within the physical layer.

4.5 Time Consumption of DT Closed-loop Control

In the control loop shown in Fig. 4.1, we can write the total commissioning time T_{tot} as:

$$T_{\text{tot}} = T_{\text{update}} + T_{\text{sim}} + T_{\text{op}}, \tag{4.12}$$

where T_{update} is the total time consumption for updating the DT, T_{sim} is the total simulation time in the QoT tool including optimization and $\widehat{SNR}$ prediction, T_{op}

is total operation time for WSS setting. Specifically:

$$T_{\text{update}} = (\lceil N_{\text{op}}/K_{\text{update}} \rceil + 1) \cdot t_{\text{update}}, \tag{4.13}$$

where N_{op} is the total number of power adjustment steps, $\lceil\ \rceil$ is the ceiling function, and the DT is updated (through monitoring) every K_{update} power adjustment steps.

Including the initialization, there are $\lceil N_{\text{op}}/K_{\text{update}} \rceil + 1$ updates. t_{update} is the time to collect data from the physical layer to update the entire DT. Then,

$$T_{\text{sim}} = (N_{\text{op}} + 1) \cdot t_{\text{sim}}, \tag{4.14}$$

where t_{sim} is the time to run the $\widehat{SNR}$ prediction tool and run the proposed algorithm in the DT; there are $N_{\text{op}} + 1$ simulations needed for each operation and the initialization, and:

$$T_{\text{op}} = N_{\text{op}} \cdot t_{\text{WSS}}, \tag{4.15}$$

where t_{WSS} is the time needed to configure the attenuation profile of WSS, thereby adjusting the launch power.

The time complexity of Alg. 4 to generate the next step is $\mathcal{O}(\frac{1}{K_m} N_{\text{OMS}}^{K_m})$. Hence, the trade-off between K_m and computation power also needs to be considered during commissioning. For a network with N_{OMS} where OMS_n has $N_{\text{span(n)}}$ spans, the upper bound for simulation time t_{sim} in Eq. (4.14) is:

$$t_{\text{sim}} \leq \frac{1}{K_m} \Big(\sum_{n}^{N_{\text{OMS}}} N_{\text{span}}(n) \Big)^{K_m} \cdot t_{\text{sim,span}}, \tag{4.16}$$

where $t_{\text{sim,span}}$ is average simulation time per span. If any of the K_m steps results in a significant degradation of the $\widehat{SNR}$, no further simulation of this step will be performed, hence Eq. (4.16) is indeed an upper bound.

Without considering any parallel data collection for updating the DT, the update time t_{update} in Eq. (4.13) can be written as:

$$t_{\text{update}} = \sum_{n}^{N_{\text{OMS}}} t_{\text{update}}(n) = \sum_{n}^{N_{\text{OMS}}} 2t_{\text{mon}} + (N_{\text{span}}(n) + 1) t_{\text{OA}}, \tag{4.17}$$

where t_{mon} is the time to get a power profile by the monitor, and t_{OA} is the time to collect data (gain, total in/output power) from an OA. The parameters refinement technique from [7] can be used to estimate the OA gain profile and lumped losses so that power monitoring is only needed for the first and last optical amplifiers of each OMS.

If monitoring data for all OMSs is collected in parallel, Eq. (4.17) becomes:

$$t_{\text{update}} = \max_{n}(t_{\text{update}}(n) + t_{\text{delay}}(n)), \tag{4.18}$$

where t_{delay} is the communication time between the controller and NE on OMS_n.

Normally, $t_{\text{update}}(n)$ is in the order of seconds while $t_{\text{sim,span}}$ is in the order of ms, then $t_{\text{update}} \gg t_{\text{sim}}$ for $K_m = 1$ and any N_{OMS}. However, it may not be true in some scenarios if $K_m \geq 2$ with a large N_{OMS}.

For instance, consider a homogeneous network in which each OMS has the same number of spans N_{span}. Then, for an upper bound on t_{sim}:

$$t_{\text{sim}} = N_{\text{OMS}}^{K_m-1} N_{\text{span}}^{K_m} \cdot t_{\text{sim,span}}, \tag{4.19}$$

$$t_{\text{update}} = N_{\text{OMS}}(2t_{\text{mon}} + (N_{\text{span}} + 1)t_{\text{OA}}), \tag{4.20}$$

The ratio $t_{\text{sim}}/t_{\text{update}}$ is then:

$$\frac{t_{\text{sim}}}{t_{\text{update}}} = \frac{1}{K_m} N_{\text{OMS}}^{K_m-1} N_{\text{span}}^{K_m} \cdot r_t, \tag{4.21}$$

where $r_t = \frac{t_{\text{sim,span}}}{2t_{\text{mon}}+(N_{\text{span}}+1)t_{\text{OA}}}$. Depending on the data collection time and computing resource for the DT, ratio r_t is typically smaller than 1/100. The time consumption could be very large when the network scales or we look a larger K_m ahead. When $K_m = 2$, the factor $\frac{1}{K_m} N_{\text{OMS}}^{K_m-1} N_{\text{span}}^{K_m} = \frac{1}{2} N_{\text{OMS}} N_{\text{span}}^2$ is still linear with N_{OMS}. However, if $K_m > 2$, then the factor $\frac{1}{K_m} N_{\text{OMS}}^{K_m-1} N_{\text{span}}^{K_m}$ is no longer linear (quadratic, ...) with N_{OMS}.

4.6 Experimental Setups

4.6.1 Network Elements

The testbed based on commercial products has 5 OMSs as shown in Fig. 4.4. The OMSs are heterogeneous, containing heterogeneous fiber spans and different types of amplifiers. The fiber types include G.652.D standard single model fiber (SSMF), G.654.E pure-silica-core fiber (PSCF), G.655 large effective area fiber (LEAF) and G.655 true-wave-RS fiber (TW-RS).

The system operates in the C-band, and EDFA is the only type of optical amplifier used in the network, the gain and tilt of EDFAs were pre-configured and not adjusted during the experiments. The EDFAs shown in Fig. 4.4 have various

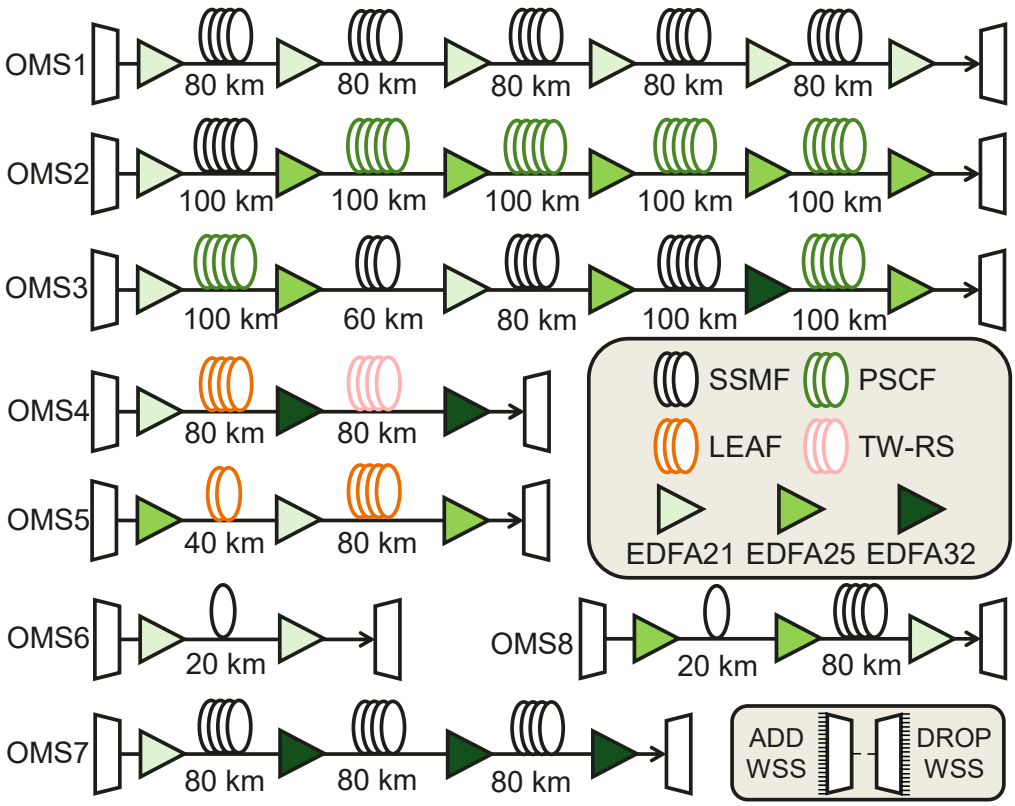


Figure 4.4: 8 OMSs in our lab. The commercial WSSs used in the testbed have 20 add/drop ports. All OMSs are fully connected and can be configured to any required topology through SDN controller.

tunable gain range in gain-lock mode, EDFA21 has a range of 16-21dB, EDFA25 has a range of 19-25dB, EDFA32 has a range of 23-32dB.

The WSS has flex grid which can set to 50/75/100/150 GHz channel spacing within the 6 THz C-band. The C-band wavelength grid of the commercial WSS is shown in the appendix B.1.

Channel loading is emulated with an ASE source or dummy light [68, 69]. Real-time commercial transponders (200 Gb/s PDM-QPSK, 400 Gb/s PDM-PCS-16QAM) are used to measure the BER, thereby calculating the SNR.

4.6.2 Network Controller and Data Collection

The testbed is controlled by an SDN controller in AI-Light presented in Section 4.2. The SDN controller collects the data from the physical layer and implements the DT to perform various optimization algorithms (routing, spectrum allocation, power optimization, ...). The power spectra can be monitored by either an optical spectrum analyzer (OSA) or a commercial optical performance monitor (OPM). Besides, we can also emulate link failures by plugging out the fibers or disabling the optical

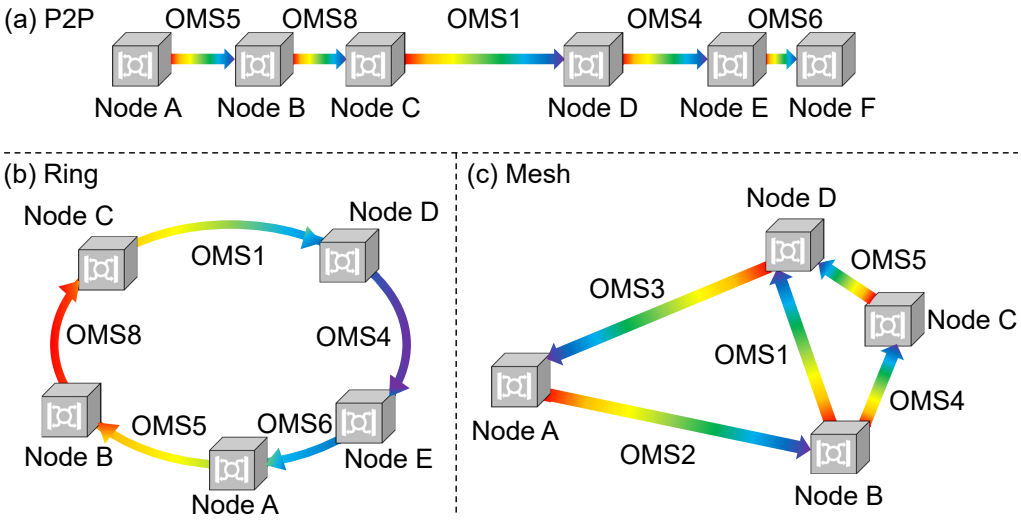


Figure 4.5: Network topologies.

amplifiers in the testbed, and collect the alarm logs from network elements.

4.6.3 Topologies

There are 3 topologies (Fig. 4.5) used in the experiments:

Topology 1: Point-to-Point [ACP2023]

A point-to-point (p2p) topology is shown in Fig. 4.5(a). We emulated 114 services (51.25% average spectrum usage) using ASE loading with 75 GHz channel spacing and a real-time 200 Gb/s PDM-QPSK transponder is used to replace the ASE loading during the SNR measurements. The services were randomly loaded according to the demand table:

Table 4.1: Service table for P2P.

Source/Destination	Light Path	Length [km]	N_{svc}
A/B	OMS 5	120	6
A/C	OMS 5–8	220	6
A/D	OMS 5–8–1	620	6
A/F	OMS 5–8–1–4–6	800	16
B/C	OMS 8	100	14
C/D	OMS 1	400	43
D/E	OMS 4	160	7
D/F	OMS 4–6	180	9
E/F	OMS 6	20	7

Table 4.2: Service table for ring

Source/Destination	Light Path	Length [km]	N_{svc}
C/B	OMS 1–4–6–5	700	5
B/A	OMS 8–1–4–6	680	5
A/E	OMS 5–8–1–4	780	5
D/C	OMS 4–6–5–8	400	5
C/A	OMS 1–4–6	580	5
C/E	OMS 1–4	560	5
B/E	OMS 8–1–4	660	5
D/B	OMS 4–6–5	300	5
A/D	OMS 5–8–1	620	5
B/D	OMS 8–1	500	5
E/C	OMS 6–5–8	240	5
C/D	OMS 1	400	5

Topology 2: Ring [OFC2024]

A ring topology is shown in Fig. 4.5(b). We load 60 services (60% average spectrum usage) based on Tab. 4.2 by ASE source with 100 GHz channel spacing and loaded based on "first-fit" allocation (in wavelength) with the "set and forget" strategy (each service is loaded with a target power, without re-adjusting the powers of the previous services) [13].

The span lumped losses uncertainty is assumed to be known (there are emulated with variable optical attenuator (VOA) with known values in our testbed). An optical channel monitor (OCM) is deployed at each optical amplifier (OA) for reference. To mimic realistic networks, the DT may use spectra at amplifiers co-located with WSS only; in this case, we use our parameters refinement technique to improve the accuracy of the DT.

Topology 3: Mesh Network [ECOC2024, JOCN2025-1]

A mesh network topology is shown in Fig. 4.5(c). We emulated 95 services based on demand Tab. 4.3 using ASE loading with 100 GHz channel spacing in the network and emulated set-and-forget loading (i.e., the power of a service may drift as further services are established,) such that channels are not well equalized.

We emulated launch power in this experiment following [13] as a Gaussian distribution with 0dB mean and 1dB standard deviation. A real-time 400 Gb/s (PDM-PCS16QAM) transponder is used to replace the ASE loading during the SNR measurements. For experimental assessment of the SNR margin of the network, we use this transponder to measure the pre-FEC BER of the 5 worst services as reported

Table 4.3: Service table for mesh network

Source/Destination	Light Path	Length [km]	N_{svc}
C/B	OMS 5–3–2	960	15
D/C	OMS 3–2–4	1100	15
D/B	OMS 3–2	940	5
A/D	OMS 2–1	900	10
B/A	OMS 1–3	840	10
A/C	OMS 2–4	660	5
C/A	OMS 5–3	560	5
B/D	OMS 1	400	10
C/D	OMS 5	120	10
B/C	OMS 4	160	10

by the DT and convert into SNR. All services have the same FEC limit SNR_{FEC}, therefore, the network-wide SNR margin simplifies to $\min_s SNR(s) - SNR_{\text{FEC}}$ (from Eq. (3.10)).

The DT needs periodic updates due to the OA gain profile variation caused by the launch power profile variation. We do not collect the gain profiles directly since most operators deploy per-channel power monitoring only at the output of booster and pre-amplifier. We apply the parameters refinement technique to refine the OA gain profile (as well as connector losses, which also cannot be directly measured) as in [7].

During each update cycle, the DT collects real-time data from the physical network, including power spectra at the output of boosters and pre-amplifiers, total input/output power of all OAs, as well as the configured values for gain and tilt of all OAs. The power spectra is monitored by an optical spectrum analyzer in the testbed.

4.7 Experimental Results

4.7.1 SNR Degradation without DTs

We demonstrate that open-loop and sequential setting with one-shot could significantly degrade some services' SNR in point-to-point [28], ring [27] networks and we also observe such a SNR degradation in the mesh network [26] due to the power propagation.

For a 5-OMS network, there are 5! = 120 possible one-shot OMS adjustment sequences. From the simulation results shown in Fig. 4.6(top), we observe that regardless of the adjustment sequence, the SNR deteriorates compared to the initial

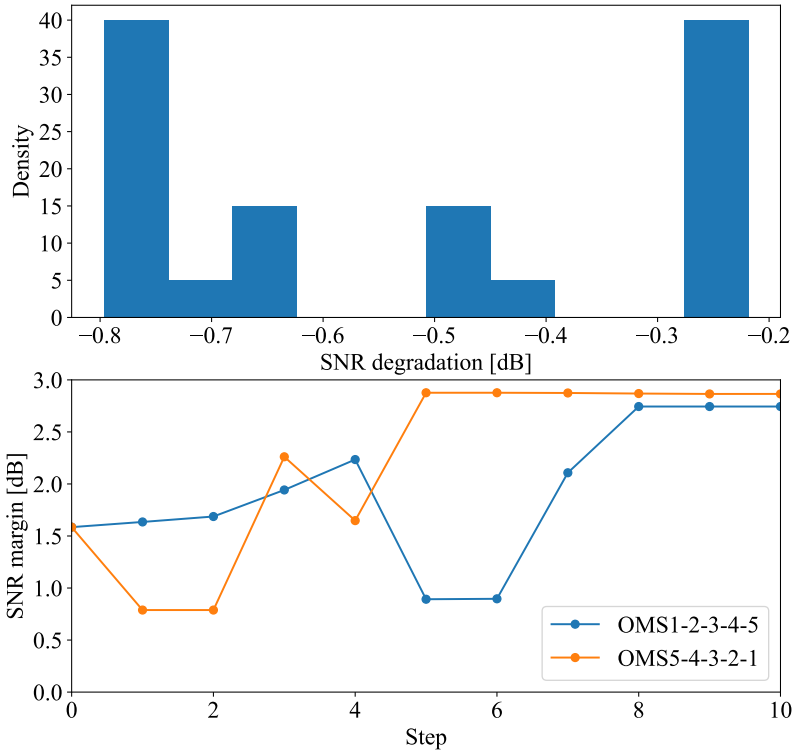


Figure 4.6: SNR degradation with open-loop. Top: Simulation results of SNR degradation value with 120 different orders. Bottom: SNR margin variation with different operation orders.

value during the process. Fig. 4.6(bottom) illustrates two of these cases. In addition to the SNR degradation, it is noteworthy that nearly 10 one-shot adjustments are required to achieve the target, optimized state in the 5-OMS network. This is primarily due to power propagation: although the power of one OMS is equalized in a single step, adjustments to the power of other OMSs cause power propagation, leading to deviations from the optimized state in previously equalized OMSs, which then need to be re-equalized.

4.7.2 Open-loop vs. Closed-loop in the Ring

In this section, we will demonstrate the experimental results based on the ring topology and show the performance of DT with different scenarios. Fig. 4.7 shows the network-wide margin for the different scenarios.

We first use the ground truth, which collects all available power spectra (per span) to assess the validity of our method. Fig. 4.7(a) shows that with the baseline method, SNR first decreases during re-optimization: the margin drops by 0.7dB to 0.3dB, which is very close to the FEC limit.

Fig. 4.7(b) shows that the DT-enabled algorithm addresses the issue: the SNR

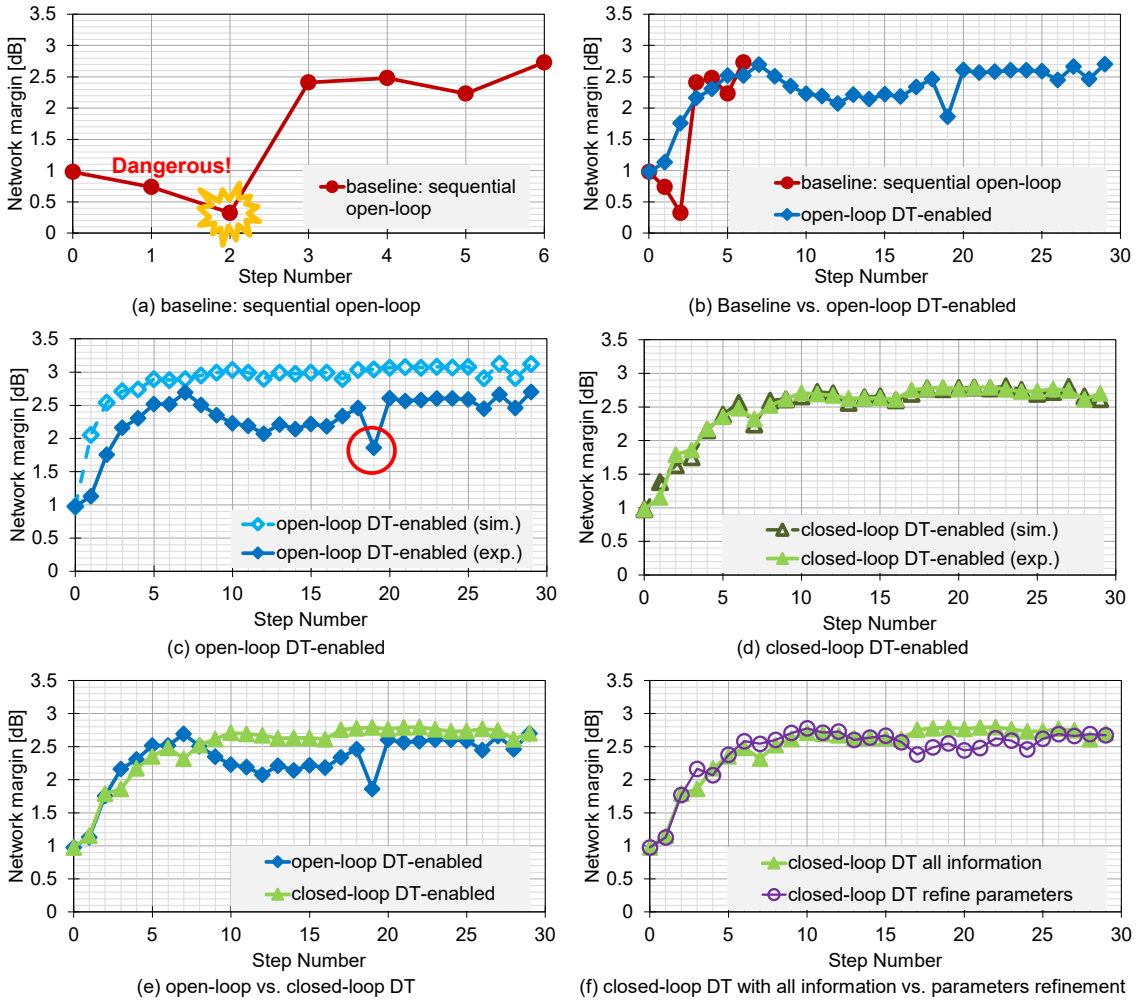


Figure 4.7: Experimental results in a ring topology network.

margin never drops below the initial value during re-optimization. Note that, as the optimization is performed network-wide, re-optimization improves not only the network SNR margin (SNR of the worst channel network-wide), but also other channels' SNR.

The gain profile of each OA is varying during the power re-configuration; this effect is not considered during optimization by the open-loop algorithm. Consequently, SNR prediction becomes slightly inaccurate with time, as shown in Fig. 4.7(c) where (sim.) denotes the DT prediction and (exp.) the monitored (ground truth) value. This prediction inaccuracy causes the observed SNR drop (Fig. 4.7(c), red circle) where the power of the service with worst SNR is different from prediction, leading to a wrong optimization choice in the DT and the SNR drop; Still, the SNR margin is much higher than the FEC limit.

By closing the loop, DT can update the input parameters in real-time to improve SNR estimation accuracy, as shown in Fig. 4.7(d).

Though both open and closed loop DT-enabled power re-optimization improve the network SNR margin by 1.5dB without interrupting the other services, the additional information gained through real-time monitoring stabilizes the SNR margin (no oscillation), see Fig. 4.7(e).

Per-span power spectra are usually not available in commercial networks; we now use only the monitored power spectra at first and last amplifier of each OMS and perform the parameters refinement technique described in [7] to estimate the missing information. As shown in Fig. 4.7(f), we obtain the same results either with realistic inputs supplemented with the refinement technique, or with full knowledge of all physical layer information.

4.7.3 Multi-step Lookahead and Parallel in the Mesh Network

For mesh network cases, we use Condition 1 (Eq. (4.7)) with $\varepsilon_{Perr} = 0.5$dB as the convergence criterion. Fig. 4.8 shows the variation of average power error with respect to the number of operational steps for the four cases defined above.

In this work, we only consider the scenarios with $K_m \leq 2$, our proposed methods focus on reducing the number of operations N_{op}, thereby decreasing the total time consumption T_{tot}.

For Case-1, it is evident that local rather than global optimization is achieved: OMS5 struggles to converge, resulting in a total of 31 steps. In Case-2, with multi-step prediction, the process escapes the local optimization trap and satisfies the convergence criterion for average power error in 18 steps. Compared to Case-1, Case-3 employs dynamic step-size adjustment and dynamic metric selection, achieving power equalization in just 14 steps. Although Case-4 still uses a small fixed step-size adjustment, parallel adjustment allows for faster overall power equalization, converging in only 13 steps.

From Fig. 4.8, we can summarize the distinctive characteristics of average power error reduction across the four cases: Case-1 exhibits symptoms of local optimization; Case-2 overcomes this limitation; Case-3 shows a significantly steeper error reduction slope; and Case-4 motivates parallel equalization.

As shown in Fig. 4.9, SNR_{margin} is improved by ∼1.5dB through power optimization. The plots include DT-*predicted* values (using monitoring data from before an operation; empty circle), the measured value after an operation (plain circle), and also the DT-estimated value (using monitoring data from after operation; cross). It shows good alignment between results from the DT (both a-priori prediction and a-posteriori estimate) and measurements.

The SNR_{margin} converges in only 10 steps for all strategies (Fig. 4.9), however,

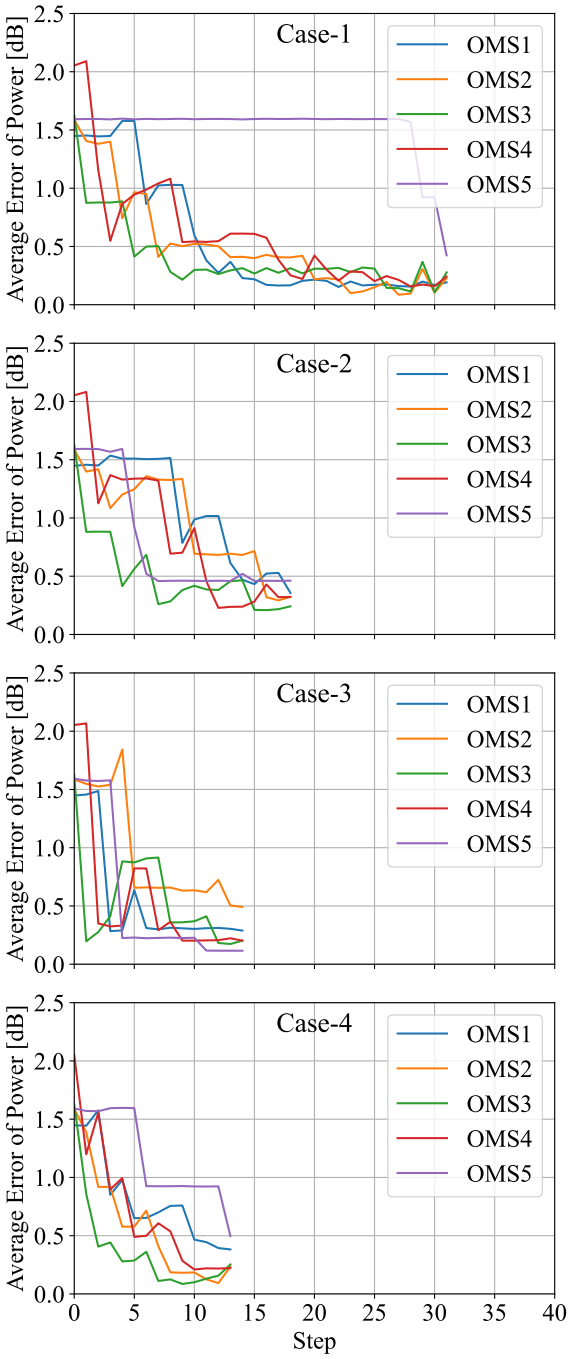


Figure 4.8: Average error of between current channel power and target channel power with steps. Case-1: 1-step lookahead with 1dB step-size; Case-2: 2-step lookahead with 1dB step-size; Case-3: 1-step lookahead with dynamic step-size; Case-4: 2-step lookahead with 1dB step-size and parallel setting.

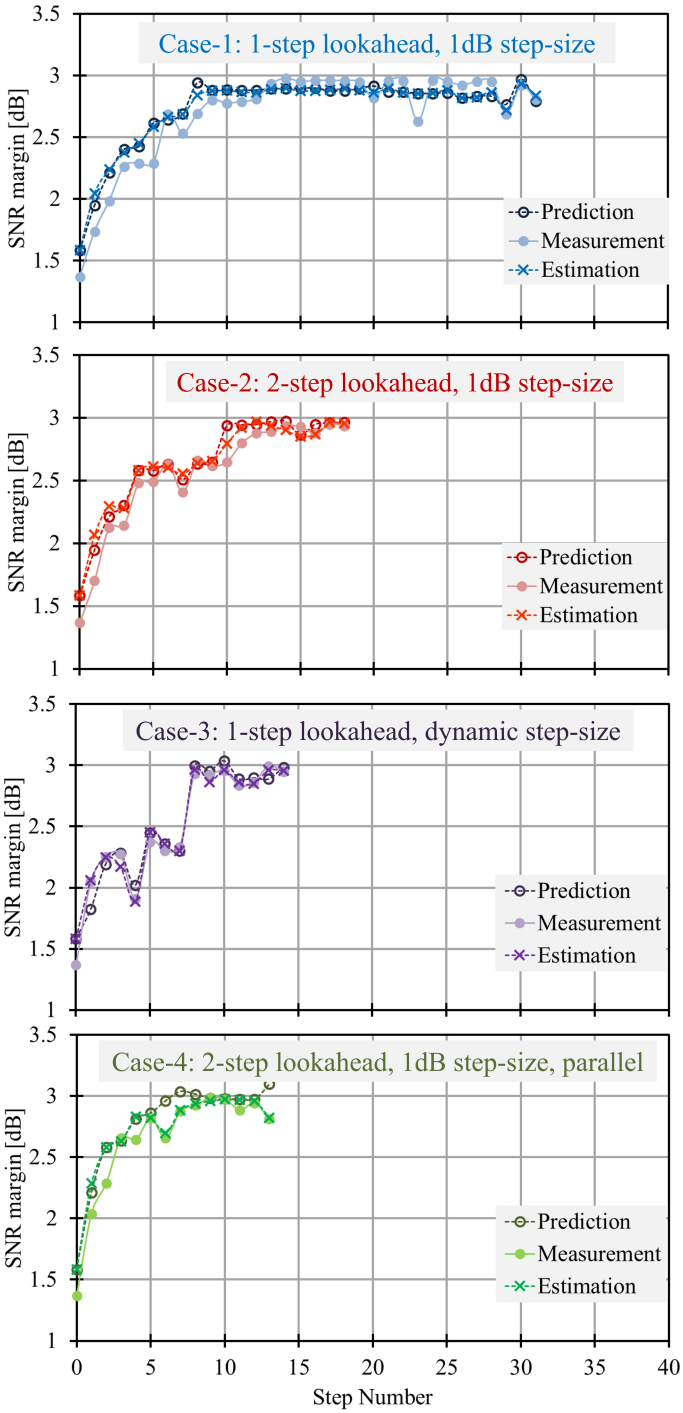


Figure 4.9: SNR margin with different scenarios. Case-1: 1-step lookahead with 1dB step-size; Case-2: 2-step lookahead with 1dB step-size; Case-3: 1-step lookahead with dynamic step-size; Case-4: 2-step lookahead with 1dB step-size and parallel setting.

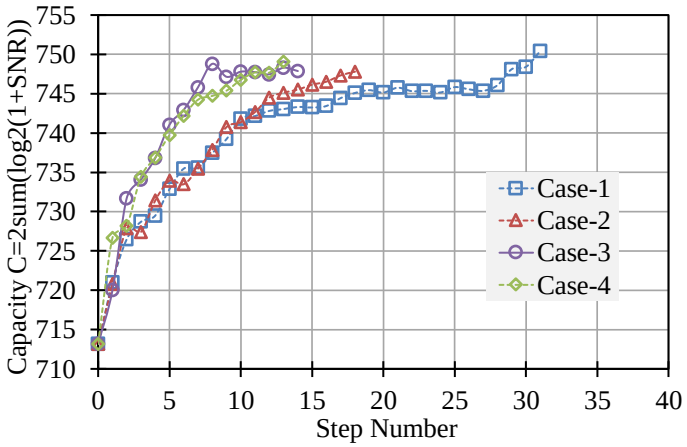


Figure 4.10: System capacity, or overall SNR. Case-1: 1-step lookahead with 1dB step-size; Case-2: 2-step lookahead with 1dB step-size; Case-3: 1-step lookahead with dynamic step-size; Case-4: 2-step lookahead with 1dB step-size and parallel setting.

total capacity converges more slowly (Fig. 4.10), which consistent with the average power error evolution (Fig. 4.8).

4.7.4 Time Consumption Analysis

Then, for the mesh network cases, we compute the total commissioning time by applying Eq. (4.12)-(4.17). Data collection is not parallel in our testbed. The results are shown in Fig. 4.11, normalized to Case-1 (as baseline) commissioning time (set to 100 for convenience).

Compared with Case-1, the proposed algorithm can save 40% (Case-2) / 53% (Case-3) / 53% (Case-4) of T_{tot}. The pie chart reveals that the major cost of T_{tot} is spent on updating the DT, indicating that increasing K_{update} to reduce update times could further save time.

The quantitative comparison is shown in Tab. 4.4 and we summarize the advantages and disadvantages of different cases in Tab. 4.5. By integrating multi-step lookahead prediction, dynamic step-size adjustments and parallel configurations, our proposed methodology ensures consistent improvements in SNR margins and efficient convergence to optimal power states without service disruptions.

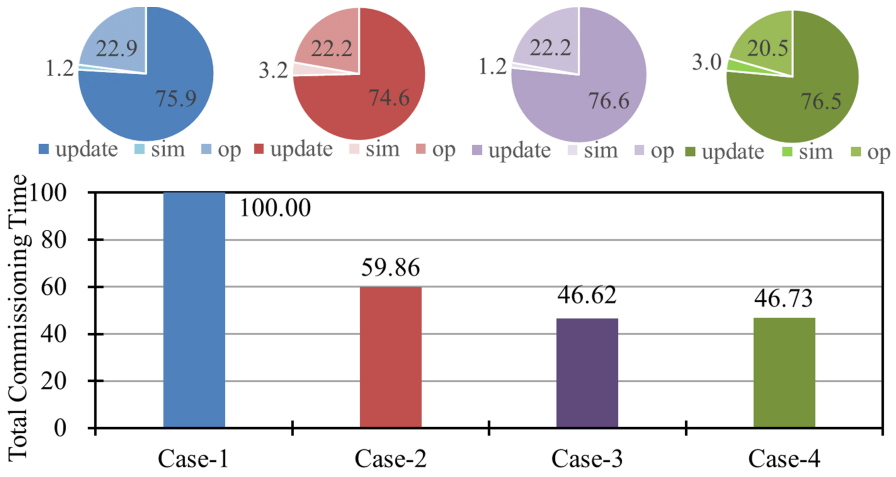


Figure 4.11: Time consumption analysis. Case-1: 1-step lookahead with 1dB step-size; Case-2: 2-step lookahead with 1dB step-size; Case-3: 1-step lookahead with dynamic step-size; Case-4: 2-step lookahead with 1dB step-size and parallel setting.

Table 4.4: Quantitative Metrics Comparison

Case	1	2	3	4
K_m	1	2	1	2
Step-size [dB]	1	1	dynamic	1
Parallel	No	No	No	Yes
Number of steps to convergence	31	18	14	13
Total commissioning time [normalized]	100	60	47	47
Final SNR margin improvement [dB]	1.6	1.5	1.6	1.5
Final capacity improvement [%]	5.2	4.9	5.1	4.9

Table 4.5: Qualitative Evaluation

Case	Conclusion
1	Baseline
2	Avoids local optima and oscillations.
3	Faster (with larger steps) but more prone to temporarily low SNR margin.
4	Even faster during operation (parallel operation) but need to check all possible WSS change orderings in the DT.

Chapter 5

AI for Optical Networks

The increasing complexity of optical networks, driven by heterogeneous equipment, dynamic traffic, and stringent performance requirements, has made traditional model-based and rule-based approaches difficult to scale. In this context, machine learning (ML), a subset of AI, has emerged as a key enabler for data-driven network operation and automation. This chapter builds upon publications produced during the course of the PhD research project [29–31].

5.1 Machine Learning and Deep Learning

5.1.1 Traditional Machine Learning

ML is a subset of AI that focuses on building systems capable of learning from data and making data-driven decisions. Unlike traditional programming, where explicit rules are manually defined, ML algorithms progressively improve their performance on a specific task by learning patterns and relationships from data. This process typically involves training a model on a dataset, during which the model adjusts its parameters to minimize prediction errors and optimize performance. ML has been widely applied in areas such as natural language processing, image recognition, recommendation systems, and predictive analytics. The ability of ML models to generalize from training data to previously unseen data makes them powerful tools for solving complex problems across various domains.

5.1.2 Neural Networks and Deep Learning

Neural networks and deep learning [70] are advanced techniques within the field of ML. Neural networks are computational models inspired by the structure and func-

tion of the human brain. They consist of layers of interconnected nodes, or neurons, that process information and pass it through the network. Each neuron applies a nonlinear transformation to its inputs and passes the result to the next layer. Deep learning refers to neural networks with many layers, allowing them to learn and represent complex patterns and hierarchies of features in data. This capability has led to significant advancements in areas such as computer vision, speech recognition, and natural language processing. Deep learning models, such as convolutional neural networks (CNNs) [71], recurrent neural networks (RNNs) [72], long short-term memory networks (LSTMs) [73], generative adversarial networks (GANs) [74], Transformers [75], autoencoders [76], variational autoencoders (VAEs) [77], graph neural netwokrs (GNNs) [78], and multilayer perceptrons (MLPs) [79] have achieved state-of-the-art performance on many tasks, often surpassing human-level accuracy. The success of deep learning is largely due to the availability of large datasets and the computational power of modern hardware, such as GPUs.

5.1.3 Applications of ML in Optical Networks

Early works such as [80–82] highlight the large volumes of heterogeneous data generated in optical networks, including performance monitoring, telemetry, alarms, and logs, which can be leveraged to improve network efficiency, reliability, and adaptability. ML techniques enable the extraction of knowledge from these data to support tasks such as modeling, failure management, traffic prediction and resource allocation, as shown in Table 5.1.

Despite these advances, current ML approaches are typically task-specific and rely primarily on structured numerical data. However, as emphasized in [80], a significant portion of network data is unstructured, including logs, configuration files, and operational documentation. This limitation motivates the exploration of large language models (LLMs) as a complementary approach. Although LLMs have not yet been widely adopted in optical networks, LLMs offer capabilities that are particularly relevant for network operation, including:

- Processing and reasoning over textual data (e.g., logs and manuals);

Table 5.1: State-of-the-art ML applications in optical networks.

Ref.	Applications
[2, 83–85]	Modeling of components, e.g. EDFA
[1–6]	QoT estimation
[86–88]	Failure management
[89–92]	Traffic prediction & resource allocation

Table 5.2: State-of-the-Art applications of LLMs and DTs in optical networks.

Ref.	Tech.	Applications
[93, 94]	LLM	Alarm analysis using cloud LLMs.
[95]	LLM	Network assistant using cloud LLMs.
[96]	LLM	Log analysis using local LLMs.
[97–99]	LLM+DT	Intent-based network using cloud LLMs.
[100]	LLM+DT	Intent-based network using local LLMs.
[101, 102]	DT	API-based network control automation.
[8, 59–62]	DT	Digital twin optical networks (DTONs).
[103]	DT	Failure management.

- Enabling natural-language interaction and intent-based networking;
- Bridging the gap between analytics outputs and operational decisions.

These capabilities align with the broader vision of autonomous optical networks and suggest that LLMs could play a key role in future network management frameworks, complementing existing ML-based and non-ML-based techniques. Researchers in optical communications and networking are currently exploring innovative applications of LLMs and DTs.

As shown in Tab. 5.2, LLMs can assist in analyzing alarm logs and troubleshooting network issues [93–96] and they can automate network management tasks by generating network configurations based on application programming interfaces (APIs) [97–102]. DTs have recently received increasing attention in optical networks for automation and management [59–62]. A network DT can analyze physical-layer monitoring data based on physical models, ML models, or hybrid approaches [8]). Moreover, DTs can be applied to failure management in optical networks [103].

Furthermore, the DT can be used as a sandbox to test proposed physical layer changes before implementing any operations in the physical world by first emulating such changes and predicting their impact in the DT itself.

However, previous works [93–95, 97–99] were using LLMs on the public cloud, leading to data privacy concerns. Though [100] used a local LLM for network auto-configurations, it only achieved 80% accuracy in translating human intents to the appropriate network configurations.

5.2 Large Language Models

LLMs are a class of AI systems that employ advanced neural network architectures, such as transformers, to process, understand, and generate natural language text. These models underpin a wide range of applications, including chatbots, text

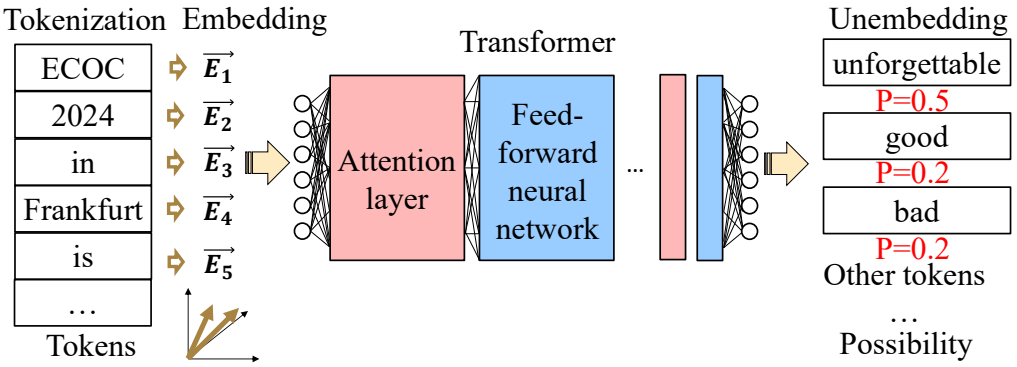


Figure 5.1: Transformer-based LLM.

summarization, language translation, and more. In particular, the Transformer architecture has revolutionized natural language processing by introducing a mechanism called self-attention, which evaluates the relative importance of words within a sentence. This mechanism enables LLMs to grasp context and relationships in text more effectively than previous methods [75].

An LLM comprises two key components: the tokenizer and the model weights stored in a safetensor file. The tokenizer converts text inputs into discrete tokens (e.g., words, subwords, or punctuation) and maps them into a high-dimensional vector space. The safetensor file encapsulates the model's learned parameters, including the weights of attention layers and feedforward neural networks. The model processes tokens through the transformer and predicts the next token with associated probabilities, enabling natural language generation.

Fig. 5.1 offers a simplified overview of the Transformer architecture, emphasizing the core components that are pertinent to its application in optical network management. This depiction is a basic representation, and actual Transformer implementations can vary significantly, incorporating different encoder and decoder blocks and other modifications tailored to specific large language models. This simplified version is intended to provide clarity and focus on the essential aspects relevant to the discussions in this paper.

Due to the complexity of transformers, these safetensor files can be exceptionally large, as they often include billions of parameters in high-precision formats like 32-bit floating-point (fp32). To address memory and hardware constraints, quantization methods reduce the precision of these parameters, enabling storage in formats such as 16-bit floating-point (fp16), 16-bit brain floating-point (bf16), or even 8-bit integers (int8). For example, an LLM with 7B parameters in fp16 has a size of $7 \cdot 10^9 \cdot 2\ Byte = 14\ GB$, while the same model with parameters in int8 only has half size, i.e. $7\ GB$, which fits on a customer-grade laptop without a pow-

erful graphics processing unit (GPU). However, quantization introduces a trade-off between memory efficiency and model accuracy, as reduced precision can degrade performance.

Pre-trained LLMs like OpenAI's ChatGPT [104], Meta's LLaMA [105], and Mistral AI [106] are built on general-purpose knowledge bases. While they contain foundational knowledge, their performance in specialized tasks remains limited due to the lack of domain-specific training.

5.2.1 Prompt Engineering and Retrieval-Augmented Generation

Prompt engineering is the practice of crafting effective input prompts to optimize the performance of LLMs. By carefully designing the phrasing, structure, and context of prompts, users can guide the model to generate more accurate, relevant, and context-aware outputs. This approach is particularly useful for tailoring LLM responses to specific tasks or domains, such as technical problem-solving or creative writing. Prompt engineering often involves iterative refinement, leveraging strategies like providing explicit instructions, examples, or formatting guidelines to maximize the model's understanding and utility.

As shown in Fig. 5.2(top), retrieval-augmented generation (RAG) is a technique that combines retrieval-based methods with generative models to enhance the accuracy and relevance of generated responses by incorporating external knowledge. The retriever searches external knowledge bases (e.g., technical documents or databases) for relevant data, which is then combined with the user query. The generator (LLM) uses this context to produce fluent, domain-specific responses. This approach ensures up-to-date, accurate, and specialized knowledge integration, especially in dynamic fields like optical networks. RAG's key benefits include im-

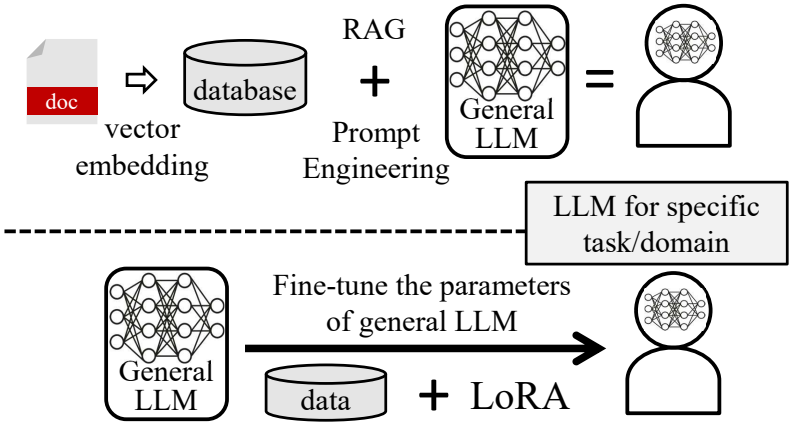


Figure 5.2: Two application-specific LLMs, using one or several of: Prompt engineering, RAG, and fine-tuning.

proved factual accuracy, adaptability to specific domains, and reduced dependency on static pre-trained knowledge.

5.2.2 Fine-Tuning the Parameters of an LLM

Since transformers are based on neural networks, another approach for adapting LLMs to specific tasks/domains is fine-tuning the parameters of the model. As shown in Fig. 5.2(bottom), the process involves further training a pre-trained model on a smaller, task-specific dataset, allowing it to specialize while leveraging previously acquired general knowledge. There are various fine-tuning approaches, including full fine-tuning, which updates all model parameters, and parameter-efficient methods like low-rank adaptation (LoRA) [107], adapter tuning, and prefix tuning, which modify only a subset of parameters or introduce new trainable elements to reduce computational costs.

LoRA freezes the pre-trained model parameters and injects trainable rank decomposition matrices into each layer of the Transformer architecture, which reduces hardware requirements and allows for efficient adaptation to new tasks.

In summary, fine-tuning integrates knowledge into the model's parameters during training, while RAG retrieves and uses external knowledge on-the-fly during inference.

5.2.3 LLM-Based AI Agents

An AI agent is an entity powered by AI, designed to perceive its environment, process data, make decisions, and take actions to achieve specific goals. These agents can learn, adapt, and interact with humans or systems, making them versatile tools across various fields. An LLM-based AI agent extends this concept by integrating an LLM with specialized tools and domain knowledge for precise, domain-specific tasks. In optical communications, AI agents can support autonomous network operation by leveraging external tools such as DTs, which are introduced in the next section. Section 5.3 further presents the design and implementation of AI agents for optical network management.

5.3 Deployment of Local AI Agents on the Testbed based on Commercial Product

Our proposed framework combines prompt engineering, RAG, and fine-tuning to address diverse tasks during the lifecycle of optical networks. To handle specialized tasks effectively, we implement a multi-agent framework, where each sub-agent has a specific role. These agents collaborate and coordinate to solve complex tasks, leveraging technologies such as LLMs, retrieval systems, and network DTs.

In the deployment of AI agents for controlling optical networks, we used the same commercial product testbed introduced in Section 4.6.

5.3.1 Local Deployed LLMs

We leverage an 8-GPU (NVIDIA Tesla V100-PCIe-32GB, launched in 2018) server for locally deploying and fine-tuning the LLMs. Each GPU has 32 GB memory and delivers 14 TFLOPS (Tera floating point operations per second) of fp32 performance. The 8-GPU memory is 256 GB, hence, we can deploy and fine-tune different open-source LLMs locally, namely Mistral-7B-Instruct [108] (model size 14.5 GB) that fits on a single GPU, Mixtral-8x7B-Instruct [109] (model size 90.4 GB), which requires multiple GPUs. These open-source models are downloaded from Hugging Face [110], the biggest community of open-source LLMs. Mistral AI develops LLMs supporting multiple European languages, and we use it to work with not only English but also French, Italian, Spanish, German, Greek, etc.

5.3.2 Local Deployed AI Agents for Control and Management of Optical Network

We build two task-specific AI agents based on LLMs and different tools: AI-Agent A1 for network automation, which feed LLMs with API descriptions for interaction with our SDN and DT; AI-Agent A2 for network management, which uses LLMs and embeds product documentation for system design and log analysis. Using multiple specialized AI agents instead of a single general-purpose agent allows each agent to be optimized for specific tasks, leading to more accurate and reliable performance. This approach reduces complexity, improves scalability, and minimizes the risk of hallucinations by leveraging task-specific expertise and focused training.

API-Calling to Control Optical Networks

As shown in Fig. 5.3, we deploy and fine-tune an LLM on the local GPU server and deploy the DT-based SDN controller on the same server to build the AI-Agent A1 for intent-based optical network automation. The SDN controller provides different JSON format APIs written in Python. The LLM is fine-tuned based on API-calling chat prompts and can generate valid predefined APIs in JSON format (compliance to standards is possible but is out of the scope of this paper). Hence, the AI-Agent A1 leverages such an LLM to translate human language requests into valid APIs with correct arguments so that SDN controller can implement them to realize the user's intent.

There are five APIs used in controlling, as shown in Tab. 5.3. The API functions utilize a structured set of arguments to define their inputs and operations. The arguments include:

source and destination: Represent the names of the source and destination nodes in the network. These are required to specify the endpoints for the service and are denoted by string values such as "A" or "B".

path: Specifies the light path for the service. It can be explicitly provided (e.g., "OMS1-OMS2"); if the path is not given explicitly, then it is determined dynamically by a routing and wavelength assignment (RWA) algorithm based on the shortest path of the light path length.

frequency: Indicates the preferred frequency for the service in MHz. For instance, 196625000 [MHz]. If not specified, the RWA algorithm selects a frequency

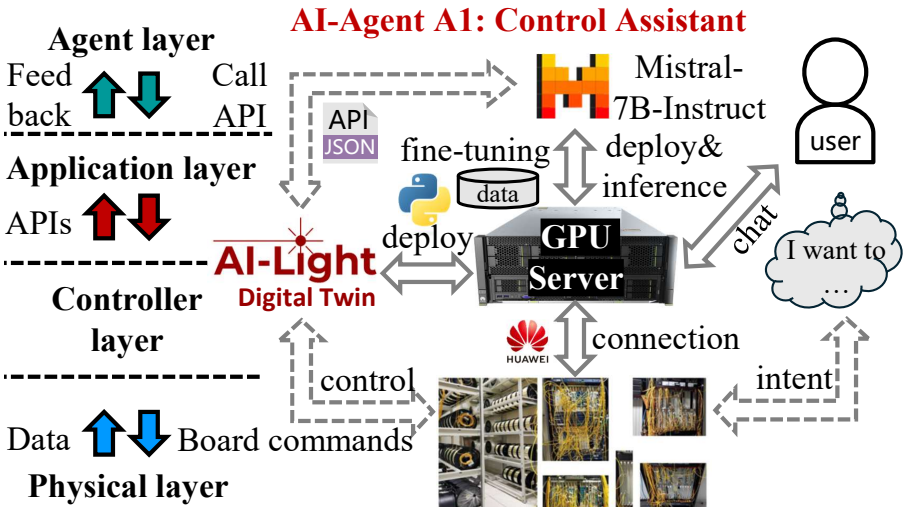


Figure 5.3: AI-Agent A1 for network control based on user's intent. (Solid arrows represent direct connections, while dash arrows represent indirect connections.)

Table 5.3: APIs used in experiments and prompts for fine-tuning AI-Agent A1.

API name	Usage	Arguments
add_och	Add a service in the network.	source, destination, path, frequency, bandwidth, board_mode
del_och	Delete a service in the network.	source, destination, frequency
estimate_qot	Estimate QoT of a service/all services.	oms_id, service_id, metric
measure_qot	Measure BER or SNR.	service_id, metric
equalize	Equalize power in OMS or network.	oms_id, method

using a "first-fit" approach, starting from longer wavelengths to shorter wavelengths.

bandwidth: Defines the bandwidth required for the service in MHz, correlating to the board mode. For instance, a 400 Gb/s service requires 100000 [MHz], while a 200 Gb/s service requires 75000 [MHz].

board_mode: Specifies the modulation format of the board. Common options include "PCS-16QAM" for a 400 Gb/s service, "QPSK" for a 200 Gb/s service, or "auto" to dynamically select an available board.

oms_id: Identifies the OMS involved in certain operations, such as QoT estimation or power equalization. Examples include "OMS1" or "ALL" to refer to all OMSs.

service_id: Acts as an index to identify a specific service in the network for operations like QoT estimation or measurement.

metric: Refers to the parameter being evaluated or optimized, such as "OSNR", "GSNR", "SNR", "ASENL" (ASE to NL Noise Ratio), or "BER".

method: Describes the approach for operations like power equalization. Common methods include "ASENL" to optimize the SNR by balancing ASE to nonlinear noise ratio, "OSNR" for equalizing OSNR, and "GSNR" for equalizing GSNR.

AI-Agent A1 should call the correct APIs with as high accuracy as possible. However, as discussed in Section 5.2, there is a trade-off between performance and general LLM size, which relates to computation resource requirements. For instance, quantization of the parameters is a standard strategy for reducing the model's size and inference costs. Ref [100] achieved 80% accuracy by using Mixtral-8x7B-Instruct with 4-bit quantization (Q4). Based on PoliMi's dataset (limited to 50

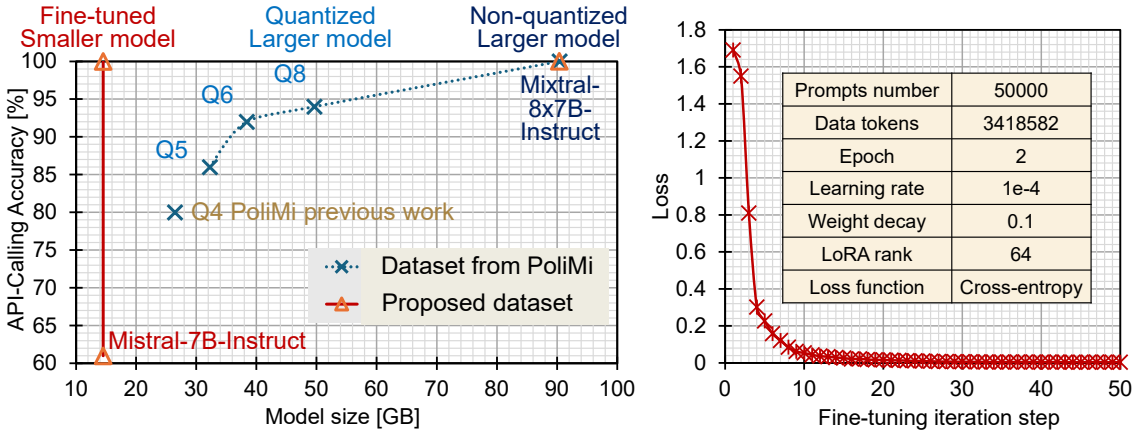


Figure 5.4: API-calling accuracy with different LLMs (left) and fine-tuning loss curve (right).

queries) made available in the same paper [100], we test the model accuracy with different levels of quantization (Q5-Q8: 5-bit to 8-bit quantization). Fig. 5.4(left) depicts the obtained model size/accuracy trade-off in blue. The results align with intuition, as larger models tend to perform better in terms of accuracy. While [100] (which leverages formal grammars) and fine-tuning both enable standard-compliant outputs, fine-tuning additionally achieves higher accuracy than [100].

We perform LoRA fine-tuning on Mistral-7B-Instruct for the five APIs we introduced in Section 5.3.2. As shown in Fig. 5.4(left), larger models achieving 100% accuracy can generate diverse data to fine-tune smaller models. For example, we create a few question-answer pairs and use the larger model to expand them with high diversity. For non-confidential data, public cloud LLMs like ChatGPT can be utilized, while private cloud LLMs are employed to augment sensitive data securely.

We generate an augmented dataset of 10k/API for fine-tuning; note that the LLM used to generate the dataset is completely independent of the LLM used in our AI agents, to avoid introducing any bias in the evaluation. The loss curve is shown in Fig. 5.4(right). We reduce the model size by 83%, from 90.4 GB to 14.5 GB, while maintaining 100% accuracy in API calls, as shown in Fig. 5.4(left,red).

In addition, another advantage of LoRA fine-tuning is that it only modifies a small portion of the parameters (here, 2.9%), therefore, the model keeps its general abilities, for instance, the multi-language ability. The Mistral AI models support several European languages and still work with requests in different languages although the training set in English.

As shown in Fig. 5.5(top), we use different languages (English, French, German, Spanish, Italian, Greek) to request different services, and the fine-tuned LLM gives us correct API calls in JSON format. Moreover, the fine-tuned model is even robust

Figure 5.5: Multi-language API-calling test. Top: Asking a fine-tuned LLM to generate JSON format API configuration using different languages. Bottom: Asking a confusing question to LLM without (w/o) and with (w/) fine-tuning (FT).

to confusing questions with mixing languages. For instance, we request the service with indicating frequency in French format, which uses "," instead of ".", yields the non-fine-tuned LLM misunderstands the value/unit (Fig. 5.5 (bottom)).

Documentation and Log Analysis

The LLM develops a specialized understanding of commercial optical products content by leveraging either RAG or fine-tuning. As shown in Fig. 5.6, we leverage these two techniques to build two types of AI-Agent A2 to integrate optical product documentation. To implement RAG, we use LangChain [111] to split the text

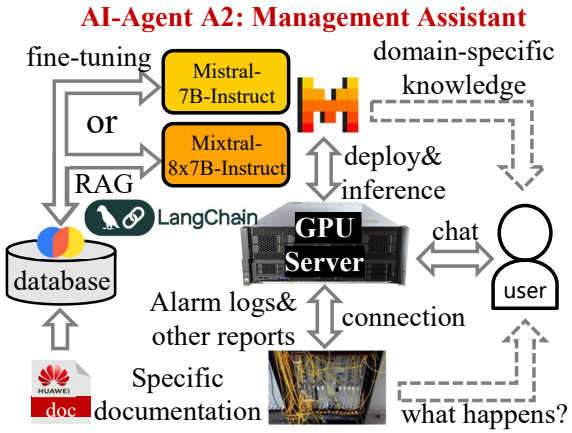


Figure 5.6: AI-Agent A2 for network management leverages knowledge from a database fed with product documentation.

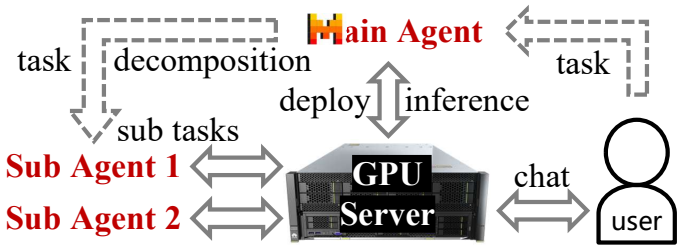


Figure 5.7: Multi-agent interaction, on the same GPU server.

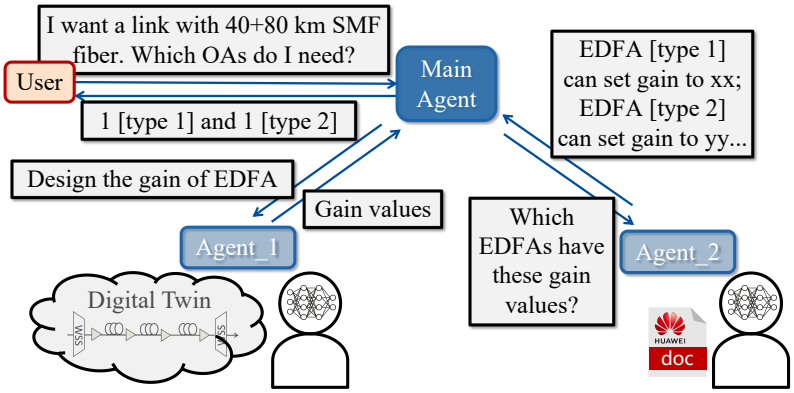


Figure 5.8: Multi-agent workflow for network design.

from documentation bodies into chunks, which are then vectorized and stored in a database with Chroma [112].

By integrating the description of logs found in product manuals, AI-Agent A2 can play a transformative role in analyzing alarm logs from optical networks, a critical aspect of maintaining network reliability and performance. Optical networks generate large alarm logs to indicate potential faults, performance degradations, or operational anomalies. Reviewing these logs is time-consuming and error-prone, especially in complex, multi-vendor environments. LLMs can automate this process by identifying patterns, classifying alarms, and providing actionable insights.

With the help of AI-Agent A2, engineers can quickly obtain troubleshooting steps, configuration guides, or product specifications tailored to their queries.

We will demonstrate RAG for product searching in Section 5.4.1, alarms management in Section 5.4.3, and fine-tuning for document integration in Section 5.4.4.

Multi-Agent

Since we deploy different AI agents on the same GPU server, we let the AI agents interact with each other to build a multi-agent system, as shown in Fig. 5.7. The main agent, which knows the roles of the other agents, is used to communicate with the other agents. An example application is network design (Fig. 5.8), which will

be discussed soon in Section 5.4.1.

5.4 Lifecycle Automation

In this section, we demonstrate how AI agents can help design/manage an optical network over its lifecycle, from Day-0 to Day-N (see Fig. 5.9).

5.4.1 Day-0: Network Design

As shown in Fig. 5.8, a multi-agent system enables collaboration between specialized agents to streamline network design.

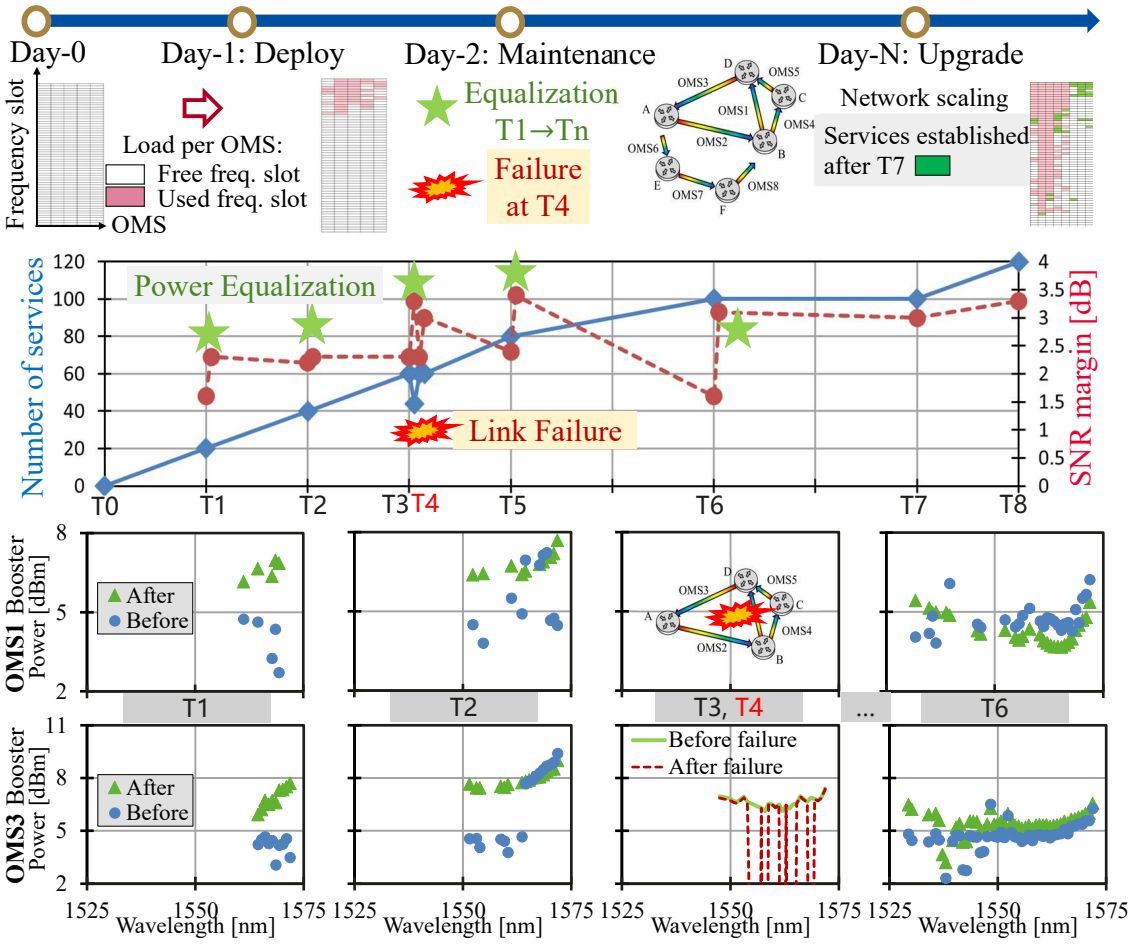


Figure 5.9: Lifecycle automation from Day-0 to Day-N. Operations and power spectra evolution in network lifecycle. Top: number of services (blue) and SNR margin evolution (red). Bottom: booster output spectra for 2 sample OMSs (OMS1 and OMS3) at different lifecycle steps. (Blue circle: before power equalization, green triangle: after power equalization. Green solid line: before link failure, red dash line: after link failure.)

A main AI-Agent is powered by prompt engineering for understanding the roles of different sub-agents. The main agent decomposes the tasks to these sub-agents when the user request requires knowledge or tools from different agents. Besides the APIs to control the optical network, we give AI-Agent A1 here the APIs to use DT for simulating and optimizing physical layer parameters, such as optical amplifier gain configuration. Once the design is complete, the main AI-Agent asks AI-Agent A2 to leverage its knowledge of commercial optical products to recommend suitable components, like the type name of amplifiers, ensuring the design is both technically optimal and implementable with available hardware. We use the JSON format messages for communication between AI agents and users, e.g., {"from": "user", "to": "Agent_Main", "content": "xxx"}, {"from": "Agent_Main", "to": "Agent1", "content": "xxx"}.

This collaborative framework bridges theoretical simulation and practical application, enhancing efficiency and feasibility in optical network development.

5.4.2 Day-1: Network Deployment

Once the optical network devices have been deployed, commissioning is required to establish the services. We start from a 5-OMS mesh network and upgrade it to 8-OMS in the end. We use AI-Agent A1 to establish the services automatically. The workflow is shown in Fig. 5.10:

Step-1: The user requests the services in natural language, for instance, "Please build a 400G service from node A to D".

Step-2: AI-Agent A1 understands the intent and generates the valid APIs sending to the DT for specific tasks introduced in section 5.3.2.

Step-3: Then the DT leverages an RWA algorithm to assign the light path

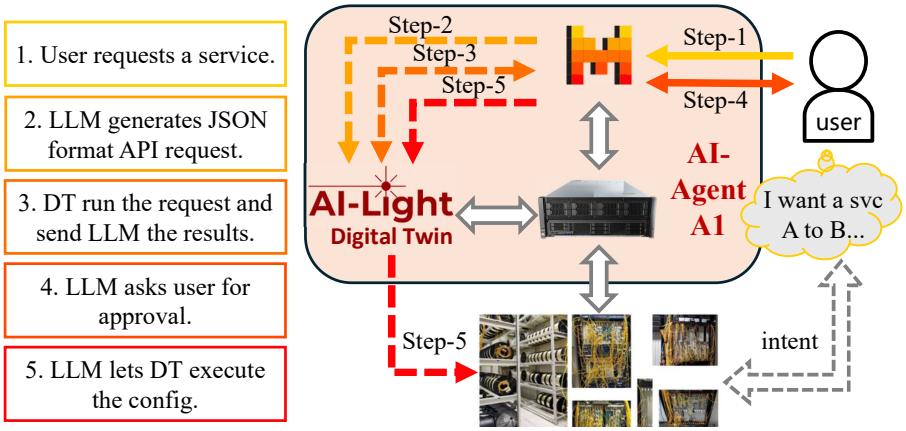


Figure 5.10: Service establishment workflow.

routing and frequency slot allocation.

Step-4: AI-Agent A1 gives suggestions to the user and requests approval.

Step-5: Upon approval, AI-Agent A1 uses the DT to configure the transmitter, receiver, and WSSs to establish the service.

5.4.3 Day-2: Network Maintenance

For network maintenance, we consider two tasks: power equalization and troubleshooting. To emulate these tasks, we load new batches of 20 services each (see slot occupation in Fig. 5.9 at T1, T2, T3, T5, T6); as introduced in Section 4.6, the services are emulated by an ASE source and we use a single transponder to replace the ASE channels during the SNR measurements ; for the following numerical evaluations, we measured SNR for 20% of the services.

Periodic Power Equalization

During the network lifecycle, amplifiers' gains vary with load, such that launch power profiles vary as well. To avoid the resulting SNR margin degradation, we use periodic service launch power re-equalization as in [13].

By chatting with AI-Agent A1, we initially add services according to the "set and forget" method, i.e., each service is established with a fixed channel power, which is not re-optimized as new services are added.

Then, we periodically re-equalize the power through API-calling by AI-Agent A1 to optimize the SNR (up to 1dB margin improvement is measured at T6) by balancing the ASE-to-NL noise ratio. For example, the power spectra before and

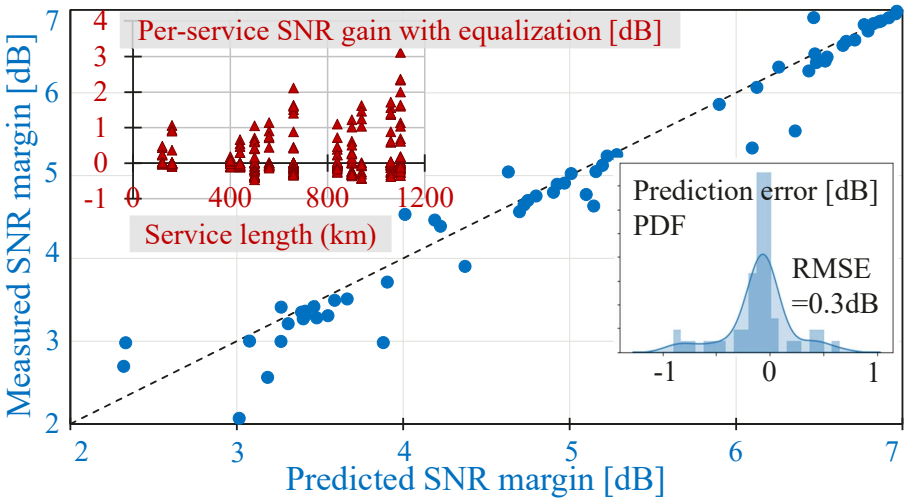


Figure 5.11: DT validation: SNR improvement with power equalization and SNR prediction accuracy.

after equalization at different times are shown in Fig. 5.9. The measured network SNR margin (defined as the minimum across all services, difference between the service SNR and its FEC limit) before and after re-equalization for each of the five batches is shown in Fig. 5.9 (jumps in the red dashed line), and the DT predicted SNR gains for all services over network lifetime in Fig. 5.11(inset). Note that some services may degrade but the network margin always improves.

The DT performs prediction and estimation of SNR before and after power re-equalization. Prediction here is the prediction of the optimized state, based on the monitoring available for the previous non-optimized network state. The probability density function (PDF) of the SNR prediction accuracy ($SNR_{prediction} - SNR_{measured}$) is shown in Fig. 5.11, the root mean square error (RMSE) is 0.3dB.

Troubleshooting

Event logs are continuously generated in optical networks, and alarms occur over time. As a natural language processing technique, LLMs are highly effective in alarm-log analysis. Based on the topology of a network and the logs from the network elements, an LLM can figure out where a failure has occurred.

The logs from the network elements contain the time stamp of each event, however, time synchronization across network elements cannot be guaranteed. As a consequence, the event time from different network elements may not be trusted for

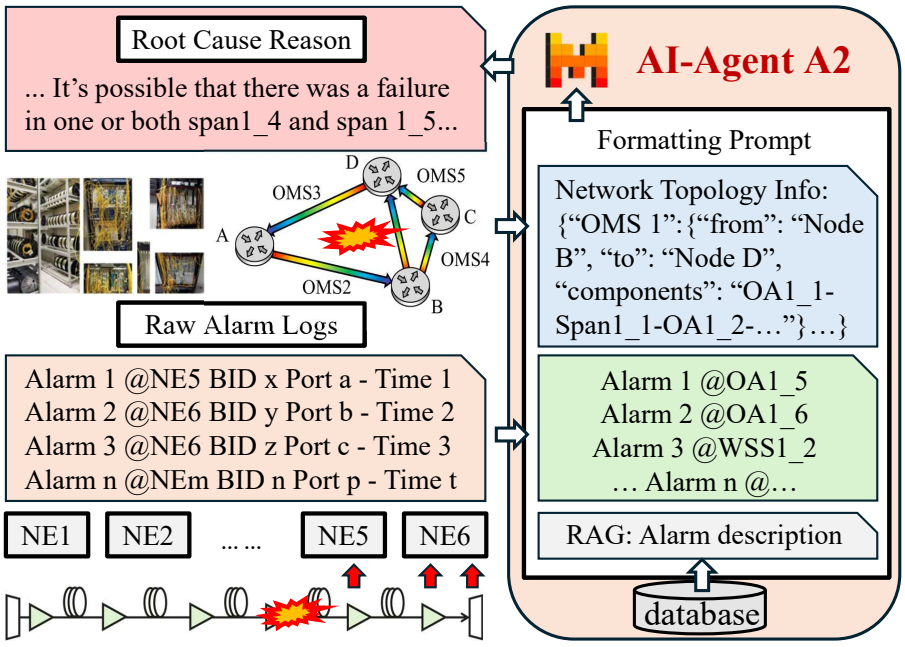


Figure 5.12: Log analysis workflow for a link failure (NE: network element; BID: board identifier; OA: optical amplifier).

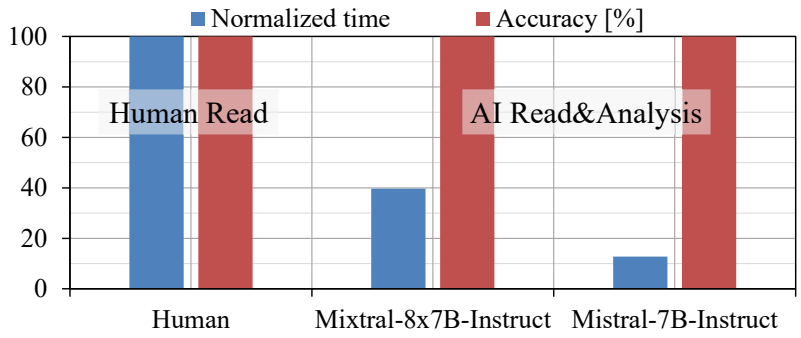


Figure 5.13: Log analysis time and accuracy comparison. Time is normalized to human reading time (as 100).

troubleshooting, and the spatial information, i.e., topology, is more important for troubleshooting. Hence, we also provide the topology information in a formatted prompt. Additionally, raw alarm logs from various network elements capture event details for individual boards, identified by their board ID and port numbers. To enhance analysis, the board ID and port are mapped to the corresponding indices of OMSs and optical amplifiers in the prompt, ensuring precise alignment with the network topology and providing clear contextual relevance. The LLM-enabled troubleshooting workflow is shown in Fig. 5.12.

We unplugged a connector to emulate a fiber cut on OMS1 (leading to the drop of 16 services) at T4 and collected the logs, which were then analyzed by AI-Agent A2. As shown in Fig. 5.13, by comparing with the reading speed of human beings [113], AI-Agent A2 performed alarm log reading and analysis 7x faster than a human would parse the logs (let alone analyze them).

In addition, we also emulate link failure by disabling in turn the amplifier located before each of the 19 spans. The LLM-based AI-Agent A2 troubleshooter successfully returns the location for each of the 19 emulated failures, as shown in Fig. 5.13. However, we need to observe that, despite different models achieving the same accuracy, we have to tune the prompt for each model. In other words, the same prompt yields different performances depending on the underlying LLM. Moreover, the order of information (topology, alarms, descriptions) also impacts the output of different LLMs.

5.4.4 Day-N: Network Upgrade

Scaling the Network

To avoid blocking at end-of-life (T7), we add 3 OMSs and use AI-Agent A1 to automate the addition of 20 services in the 8-OMS network, using the same method

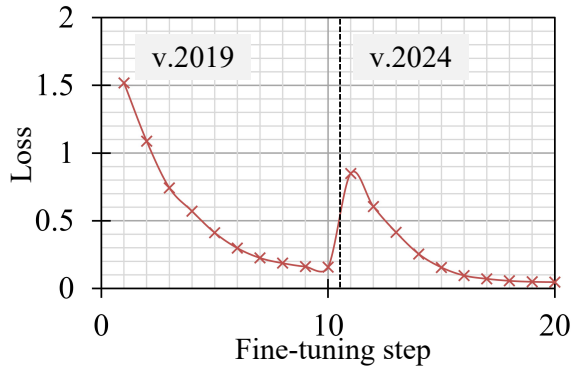


Figure 5.14: Commercial products documentation-based fine-tuning before/after upgrade. Version 2019 is for C-band commercial equipment only, while version 2024 updates L-band commercial equipment.

as in Day-2, reaching a total of 120 services established, see Fig. 5.9 (right).

Hardware and Software Upgrade

While RAG retrieves relevant information from vast technical repositories, fine-tuning adapts an LLM to deeply understand the structure, terminology, and context of domain-specific documentation. By fine-tuning the model on commercial product manuals, the LLM develops a specialized understanding of commercial optical networks content. This ensures that even without information retrieval, the model can independently provide accurate and context-aware insights.

However, different products and features are typically released during the 10-20 years lifetime of optical equipment. The AI-Agent A2 also requires to be updated. Since the safetensor is just a checkpoint of the training (cf. Section 5.2), model fine-tuning can be done iteratively on the previous model each time new product documentation or a new software version is released. Safetensors enable continuous fine-tuning using prior checkpoints.

For example, a model initially fine-tuned with documentation from 2019 was subsequently enhanced using updated 2024 documentation, resulting in improved performance while preserving earlier optimizations. The training dataset comprised specifications from various commercial products, including WSSs, EDFAs, transponders, etc. Notably, the 2019 dataset focuses exclusively on C-band equipment, whereas the 2024 dataset incorporates updated specifications for L-band equipment, reflecting technology advancements. The training loss curve (Fig. 5.14) demonstrates smooth convergence each time fine-tuning is triggered, showcasing the practicality of incremental model upgrades.

Chapter 6

Conclusions and Future Perspectives

This thesis presented a set of initial investigations toward autonomous optical network control and management, combining DT-enabled network automation with emerging AI paradigms such as LLMs and AI agents. The contributions are positioned as proof-of-concept studies that demonstrate feasibility, explore design trade-offs, and establish foundational building blocks for future autonomous optical networks. The main findings of this thesis can be summarized as follows:

- DT-based closed-loop control enables predictive and stable power optimization: the proposed multi-step lookahead strategies significantly improves convergence speed while reducing oscillatory behavior, demonstrating the value of foresight in optical network control.
- There exists a clear trade-off between performance and complexity: advanced strategies such as dynamic step sizing and parallel exploration further improve convergence efficiency, but introduce additional system complexity and transient performance degradation.
- A layered DT + AI agent SDN architecture is feasible: the proposed four-layer framework successfully integrates physical-layer modeling, control logic, and AI-driven orchestration, providing a viable blueprint for autonomous optical networks.
- LLM-based AI agents enable flexible and user-friendly network interaction: compared to traditional approaches, LLMs significantly improve adaptability, particularly for intent-based network control and management.
- LLMs are not yet suitable for critical control loops: limitations in interpretability, latency, and reliability (e.g., hallucinations) require that they be used as supervisory or assistive tools rather than fully autonomous decision-makers.

6.1 DT for Optical Networks

The proposed multi-step lookahead power equalization strategy demonstrates the importance of incorporating foresight into control decisions. Increasing the lookahead depth significantly improves convergence speed and stability by reducing local optima and oscillatory behavior. The introduction of dynamic step sizes and parallel exploration further enhances performance, achieving faster convergence with fewer iterations. However, these improvements come at the cost of increased algorithmic complexity and potential transient degradation of signal quality.

These results highlight a fundamental trade-off between convergence efficiency, robustness, and operational constraints, which is central to practical deployment in real optical networks. More broadly, this work illustrates how DT-based approaches can enable predictive and proactive control strategies that are difficult to achieve with traditional reactive methods.

Future research should address DT scalability, particularly through network partitioning and distributed simulation, to accelerate data collection and computation. Beyond control, DTs also open new application opportunities, including risk-map generation for network health assessment, QoT prediction for unseen lightpaths, and synthetic data generation for AI model training.

6.2 AI for Optical Networks

This thesis also investigated LLM-based AI agents as an enabling technology for autonomous optical network control and management. As summarized in Table 6.1, LLM-based approaches offer significantly higher flexibility and usability compared to traditional machine learning and rule-based methods. Their ability to interpret natural language and generalize across tasks reduces the need for domain-specific expertise and enables more intuitive human-network interaction.

From an application perspective (Table 6.2), LLMs show strong potential for

Table 6.1: Comparison of LLM-based, traditional ML, and rule-based approaches.

Technique	LLMs	Traditional ML	Rule-based
Flexibility	High	Moderate	Low
Interpretability	Low	Moderate	High
Scalability	High	Moderate	Low
Ease of use	High	Moderate	Low

Table 6.2: Pros and cons of applications of LLMs for optical networks control and management.

App.	Control	Management
Pros	Intent-based networking using natural language. Reduces manual intervention, minimizing human error. Speeds up network configuration.	Quickly analyzes large volumes of network logs. Reduces the need for highly specialized knowledge about vendors' hardware and expertise in software.
Cons	Requires an accurate DT. Requires high computational resources. Adds latency for time-sensitive applications.	Data privacy concerns. Risk of hallucinations (false positives or negatives). Requires continuous training and data/API governance.

intent-based control and network management. In control scenarios, they simplify configuration workflows and reduce manual intervention, while in management tasks they facilitate efficient analysis of large volumes of textual network data.

However, several limitations remain. LLM-based approaches suffer from limited interpretability, high computational requirements, and potential reliability issues such as hallucinations. In addition, their effectiveness is closely tied to the accuracy of the underlying DT models, and their latency may limit applicability in time-sensitive control loops. These constraints indicate that LLMs are not yet suitable for fully autonomous operation in critical network functions.

It is important to note that LLMs should not be viewed as the final solution for autonomous optical networks. Rather, they should be seen as an intermediate step within a rapidly evolving AI landscape. The broader concept of AI agents—capable of reasoning, planning, and interacting with network systems—remains highly relevant, and future advances in AI are expected to provide more reliable, efficient, and specialized solutions for network control and management.

Overall, this thesis demonstrates that combining DTs with LLM-based AI agents is a promising direction toward autonomous optical networks. DTs provide a reliable and physics-aware foundation for modeling and prediction, while LLMs introduce flexibility and high-level reasoning capabilities for interaction and orchestration.

At the current stage, these technologies should be viewed as complementary components that augment existing control and management frameworks rather than replace them. Achieving fully autonomous optical networks will require further advances in model accuracy, system scalability, real-time performance, and trustworthiness. This includes improving DT fidelity, developing hybrid control

architectures that combine deterministic and AI-driven methods, and establishing robust validation and governance mechanisms for AI-based decision-making.

In this context, the contributions of this thesis provide a first step toward bridging the gap between traditional optical network engineering and emerging AI-driven automation paradigms.

Appendix A

Network State in AI-Light DT

The network state used in the AI-Light digital twin is composed of several DataFrames. Fig. A.1 presents the network topology information, representing the source and destination nodes of different OMSs. Fig. A.2 shows the control ports of the connections between different network elements (OMS, ROADM, TRX, and DL). Fig. A.3 presents the line parameters DataFrame for an OMS. Fig. A.4 and Fig. A.5 show the Add-WSS and Drop-WSS DataFrames of the OMS, respectively. Fig. A.6 presents the service (optical channel, OCH) table.

Source/Destination	A	B	C	D	E
A		OMS8	OMS1		OMS6
B			OMS3		
C	OMS2			OMS5	
D	OMS4				
E				OMS7	

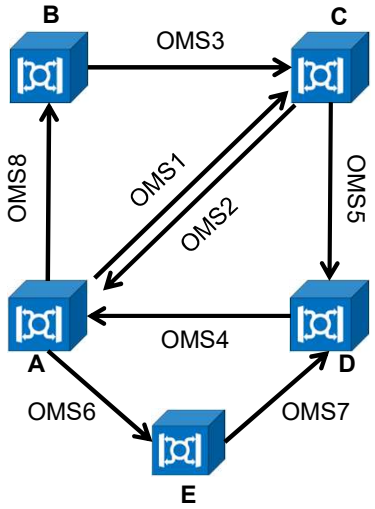
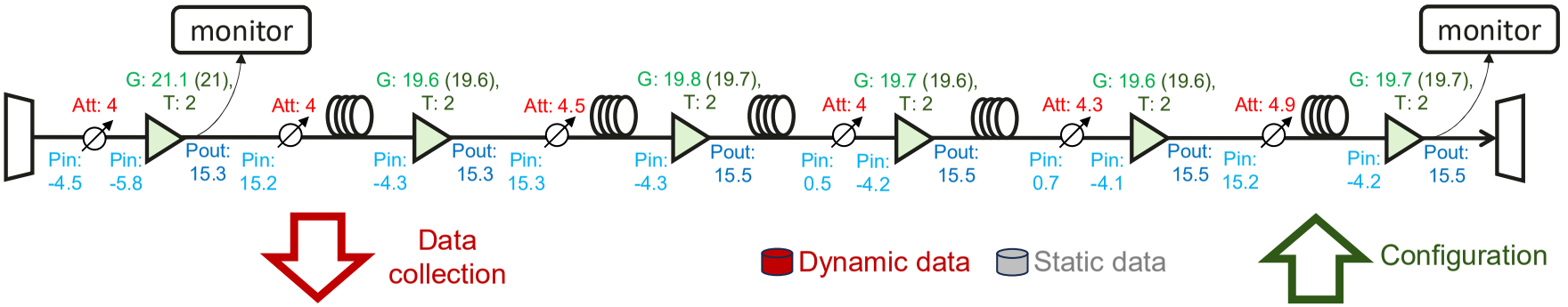


Figure A.1: Network topology information DataFrame (top) and corresponding topology (bottom).

Source/Destination	ROADM1_TX	OMS1	OMS2	OMS3	OMS4	OMS5	OMS6	OMS7	OMS8	ROADM1_RX	RX1	RX...
TX1	0-23											
TX...	0-xx											
ROADM1_TX		6-31	8-28	10-28	12-28	14-28	16-28	18-28	20-28	4-4		
OMS1			6-26	5-24	20-40	12-25	7-24	13-24	14-24	8-6		
OMS2		4-25		6-26	12-32	14-34	13-25	15-25	17-25	8-8		
OMS3		12-32	5-25		6-26	7-26	13-33	15-29	16-29	8-10		
OMS4		16-36	7-29	9-27		4-24	6-26	14-27	15-27	8-12		
OMS5		10-28	4-27	5-32	22-38		12-32	9-31	11-31	8-14		
OMS6		12-33	4-32	5-25	7-24	9-32		6-26	13-32	8-16		
OMS7		13-29	14-34	15-29	16-29	17-29	18-29		6-26	8-18		
OMS8		6-26	4-31	5-31	14-34	7-30	16-36	9-39		8-20		
ROADM1_RX											23-0	xx-0
DL		0-42	0-42	0-42	0-42	0-42	0-42	0-42	0-42			

Figure A.2: Connection information DataFrame. The source–destination port pair represents the software-level port IDs at the source and destination elements. In this example, the OMSs are fully connected, enabling flexible topology reconfiguration based on network demands.



Reference	Type	Item	NE_IP	Board_ID	Port_ID	Monitor_port	Attenuation	Gain	Tilt	Length	Pin	Pout	Mode	Gain_cfg	Span_loss
******	*********	WSS	129.9.0.40	1	2										
******	*********	VOA	***.*.**	5	2		4				-4.5				
******	*********	EDFA	***.*.**	4	1	1-2		21.1	2		-5.8	15.3	gainlock	21	
******	*********	VOA	***.*.**	4	1		4				15.2				
Span 1	SSMF	Fiber								80					19.6
******	*********	EDFA	***.*.**	6	1	1-4		19.6	2		-4.3	15.3	gainlock	19.6	
******	*********	VOA	***.*.**	4	2		4.5				15.3				
Span 2	SSMF	Fiber								80					19.6
******	*********	EDFA	***.*.**	4	2	1-6		19.8	2		-4.3	15.5	gainlock	19.7	
Span 3	SSMF	Fiber								80					19.7
******	*********	VOA	***.*.**	6	1		4				0.5				
******	*********	EDFA	***.*.**	6	2	1-8		19.7	2		-4.2	15.5	gainlock	19.6	
Span 4	SSMF	Fiber								80					19.6
******	*********	VOA	***.*.**	6	2		4.3				0.7				
******	*********	EDFA	***.*.**	5	2	1-10		19.6	2		-4.1	15.5	gainlock	19.6	
******	*********	VOA	***.*.**	3	1		4.9				15.2				
Span 5	SSMF	Fiber								80					19.7
******	*********	EDFA	***.*.**	3	1	1-12		19.7	0		-4.2	15.5	gainlock	19.7	
******	*********	WSS	***.*.**	3	1										

Figure A.3: OMS line parameters DataFrame. The units are as follows: attenuation, gain, tilt, configured gain, and span loss in dB; P_{in} and P_{out} in dBm; and fiber length in km.

port	frequency	wavelength	bandwidth	attenuation	OA0_Pout	OA5_Pout
36	196000000	1529.55	150000	0	6.36	1.84
36	194950000	1537.79	150000	5.6	6.36	6.14
42	194800000	1538.98	150000	4.9	6.26	6.24
42	193600000	1548.51	150000	4.4	6.56	8.44
42	193000000	1553.33	150000	4.5	6.36	7.84
36	191950000	1561.83	150000	4.7	6.16	6.94
42	191800000	1563.05	150000	4	6.56	7.54
36	190900000	1570.42	150000	4.3	5.46	4.44

Figure A.4: Add-WSS DataFrame. The units are as follows: frequency and bandwidth in MHz, wavelength in nm, attenuation in dB, and booster/pre-amplifier output power (P_{out}) in dBm.

port	frequency	wavelength	bandwidth	attenuation
12	193600000	1548.51	150000	0
12	191800000	1563.05	150000	0

Figure A.5: Drop-WSS DataFrame. The units are as follows: frequency and bandwidth in MHz, wavelength in nm, and attenuation in dB.

Service_ID	Frequency	Wavelength	Bandwidth	Light_Path	Transponder
1	196000000	1529.55	150000	OMS8	TX1_800G
2	194800000	1538.98	150000	OMS1	TX2_800G
3	193600000	1548.51	150000	OMS1-OMS5	TX3_800G
4	192400000	1558.17	150000	OMS6	TX4_800G
5	191200000	1567.95	150000	OMS3-OMS2	TX5_800G
...	...	...	...	...	...

Figure A.6: Service (OCH) table DataFrame. The units are as follows: frequency and bandwidth in MHz, wavelength in nm, and attenuation in dB.

Appendix B

WDM Wavelength Grid

B.1 WDM Wavelength Grid in Super C-band

Table B.1: C120 wave grid with 50 GHz channel spacing in 6 THz frequency range from 190675 GHz to 196675 GHz. (f_c and λ_c are central frequency and central wavelength of the optical channel.)

Index	f_c [GHz]	λ_c [nm]
1	196650	1524.497625
2	196600	1524.885341
3	196550	1525.273254
4	196500	1525.661364
5	196450	1526.049672
6	196400	1526.438177
7	196350	1526.826881
8	196300	1527.215782
9	196250	1527.604882
10	196200	1527.994179
11	196150	1528.383676
12	196100	1528.773371
13	196050	1529.163264
14	196000	1529.553357
15	195950	1529.943649
16	195900	1530.33414
17	195850	1530.72483
18	195800	1531.11572
19	195750	1531.50681

Index	f_c [GHz]	λ_c [nm]
20	195700	1531.898099
21	195650	1532.289589
22	195600	1532.681278
23	195550	1533.073168
24	195500	1533.465258
25	195450	1533.857549
26	195400	1534.250041
27	195350	1534.642734
28	195300	1535.035627
29	195250	1535.428722
30	195200	1535.822018
31	195150	1536.215516
32	195100	1536.609216
33	195050	1537.003117
34	195000	1537.397221
35	194950	1537.791526
36	194900	1538.186034
37	194850	1538.580744
38	194800	1538.975657
39	194750	1539.370773
40	194700	1539.766091
41	194650	1540.161613
42	194600	1540.557338
43	194550	1540.953267
44	194500	1541.349398
45	194450	1541.745734
46	194400	1542.142274
47	194350	1542.539017
48	194300	1542.935965
49	194250	1543.333117
50	194200	1543.730474
51	194150	1544.128035
52	194100	1544.525801
53	194050	1544.923772
54	194000	1545.321948
55	193950	1545.72033

Index	f_c [GHz]	λ_c [nm]
56	193900	1546.118917
57	193850	1546.51771
58	193800	1546.916708
59	193750	1547.315912
60	193700	1547.715323
61	193650	1548.114939
62	193600	1548.514762
63	193550	1548.914792
64	193500	1549.315028
65	193450	1549.715472
66	193400	1550.116122
67	193350	1550.51698
68	193300	1550.918044
69	193250	1551.319317
70	193200	1551.720797
71	193150	1552.122485
72	193100	1552.524381
73	193050	1552.926485
74	193000	1553.328798
75	192950	1553.731319
76	192900	1554.134049
77	192850	1554.536987
78	192800	1554.940135
79	192750	1555.343492
80	192700	1555.747058
81	192650	1556.150833
82	192600	1556.554818
83	192550	1556.959013
84	192500	1557.363418
85	192450	1557.768033
86	192400	1558.172859
87	192350	1558.577894
88	192300	1558.983141
89	192250	1559.388598
90	192200	1559.794266
91	192150	1560.200146

Index	f_c [GHz]	λ_c [nm]
92	192100	1560.606236
93	192050	1561.012538
94	192000	1561.419052
95	191950	1561.825778
96	191900	1562.232715
97	191850	1562.639864
98	191800	1563.047226
99	191750	1563.454801
100	191700	1563.862587
101	191650	1564.270587
102	191600	1564.6788
103	191550	1565.087225
104	191500	1565.495864
105	191450	1565.904717
106	191400	1566.313783
107	191350	1566.723062
108	191300	1567.132556
109	191250	1567.542264
110	191200	1567.952186
111	191150	1568.362323
112	191100	1568.772674
113	191050	1569.18324
114	191000	1569.594021
115	190950	1570.005017
116	190900	1570.416228
117	190850	1570.827655
118	190800	1571.239298
119	190750	1571.651156
120	190700	1572.06323

B.2 WDM Wavelength Grid in Super L-band

Table B.2: L120 wave grid with 50 GHz channel spacing in 6 THz frequency range from 184325 GHz to 190325 GHz. (f_c and λ_c are central frequency and central wavelength of the optical channel.)

Index	f_c [GHz]	λ_c [nm]
1	190300	1575.36762
2	190250	1575.781645
3	190200	1576.195889
4	190150	1576.61035
5	190100	1577.025029
6	190050	1577.439926
7	190000	1577.855042
8	189950	1578.270376
9	189900	1578.685929
10	189850	1579.101701
11	189800	1579.517692
12	189750	1579.933903
13	189700	1580.350332
14	189650	1580.766981
15	189600	1581.18385
16	189550	1581.600939
17	189500	1582.018248
18	189450	1582.435777
19	189400	1582.853527
20	189350	1583.271497
21	189300	1583.689688
22	189250	1584.1081
23	189200	1584.526734
24	189150	1584.945588
25	189100	1585.364664
26	189050	1585.783962
27	189000	1586.203481
28	188950	1586.623223
29	188900	1587.043187
30	188850	1587.463373

Index	f_c [GHz]	λ_c [nm]
31	188800	1587.883782
32	188750	1588.304413
33	188700	1588.725268
34	188650	1589.146345
35	188600	1589.567646
36	188550	1589.98917
37	188500	1590.410918
38	188450	1590.832889
39	188400	1591.255085
40	188350	1591.677505
41	188300	1592.100149
42	188250	1592.523017
43	188200	1592.946111
44	188150	1593.369429
45	188100	1593.792972
46	188050	1594.21674
47	188000	1594.640734
48	187950	1595.064953
49	187900	1595.489399
50	187850	1595.91407
51	187800	1596.338967
52	187750	1596.764091
53	187700	1597.189441
54	187650	1597.615017
55	187600	1598.040821
56	187550	1598.466852
57	187500	1598.893109
58	187450	1599.319595
59	187400	1599.746307
60	187350	1600.173248
61	187300	1600.600416
62	187250	1601.027813
63	187200	1601.455438
64	187150	1601.883291
65	187100	1602.311374
66	187050	1602.739685

Index	f_c [GHz]	λ_c [nm]
67	187000	1603.168225
68	186950	1603.596994
69	186900	1604.025993
70	186850	1604.455221
71	186800	1604.884679
72	186750	1605.314367
73	186700	1605.744285
74	186650	1606.174433
75	186600	1606.604812
76	186550	1607.035422
77	186500	1607.466263
78	186450	1607.897334
79	186400	1608.328637
80	186350	1608.760172
81	186300	1609.191938
82	186250	1609.623936
83	186200	1610.056165
84	186150	1610.488627
85	186100	1610.921322
86	186050	1611.354249
87	186000	1611.787409
88	185950	1612.220801
89	185900	1612.654427
90	185850	1613.088286
91	185800	1613.522379
92	185750	1613.956705
93	185700	1614.391265
94	185650	1614.82606
95	185600	1615.261088
96	185550	1615.696351
97	185500	1616.131849
98	185450	1616.567582
99	185400	1617.003549
100	185350	1617.439752
101	185300	1617.87619
102	185250	1618.312864

Index	f_c [GHz]	λ_c [nm]
103	185200	1618.749773
104	185150	1619.186919
105	185100	1619.6243
106	185050	1620.061918
107	185000	1620.499773
108	184950	1620.937864
109	184900	1621.376193
110	184850	1621.814758
111	184800	1622.253561
112	184750	1622.692601
113	184700	1623.131879
114	184650	1623.571395
115	184600	1624.011148
116	184550	1624.451141
117	184500	1624.891371
118	184450	1625.331841
119	184400	1625.772549
120	184350	1626.213496

Bibliography

[1] I. Sartzetakis, K. K. Christodoulopoulos, and E. M. Varvarigos, "Accurate Quality of Transmission Estimation With Machine Learning," *Journal of Optical Communications and Networking*, vol. 11, no. 3, pp. 140–150, 2019.

[2] A. Mahajan, K. Christodoulopoulos, R. Martínez, S. Spadaro, and R. M. noz, "Modeling EDFA Gain Ripple and Filter Penalties With Machine Learning for Accurate QoT Estimation," *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2616–2629, 2020.

[3] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks," *Journal of Optical Communications and Networking*, vol. 13, no. 4, pp. B60–B71, 2021.

[4] R. Ayassi, A. Triki, N. Crespi, R. Minerva, and M. Laye, "Survey on the use of machine learning for quality of transmission estimation in optical transport networks," *Journal of Lightwave Technology*, vol. 40, no. 17, pp. 5803–5815, 2022.

[5] E. Seve, J. Pesic, and Y. Pointurier, "Associating machine-learning and analytical models for quality of transmission estimation: combining the best of both worlds," *Journal of Optical Communications and Networking*, vol. 13, no. 6, pp. C21–C30, 2021.

[6] J. Müller, S. K. Patri, T. Fehenberger, H. Griesser, J.-P. Elbers, and C. Mas-Machuca, "QoT estimation using EGN-assisted machine learning for multi-period network planning," *Journal of Optical Communications and Networking*, vol. 14, no. 12, pp. 1010–1019, 2022.

[7] N. Morette, H. Hafermann, Y. Frignac, and Y. Pointurier, "Machine learning enhancement of a digital twin for wavelength division multiplexing network performance prediction leveraging quality of transmission parameter refine-

ment," *Journal of Optical Communications and Networking*, vol. 15, no. 6, pp. 333–343, 2023.

[8] Y. He, Z. Zhai, L. Dou, L. Wang, Y. Yan, C. Xie, C. Lu, and A. P. T. Lau, "Improved QoT estimations through refined signal power measurements and data-driven parameter optimizations in a disaggregated and partially loaded live production network," *Journal of Optical Communications and Networking*, vol. 15, no. 9, pp. 638–648, 2023.

[9] Y. Pointurier, "Design of low-margin optical networks," *Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. A9–A17, 2017.

[10] O. Karandin, A. Ferrari, F. Musumeci, Y. Pointurier, and M. Tornatore, "Low-margin optical-network design with multiple physical-layer parameter uncertainties," in *European Conference on Optical Communication (ECOC) 2022*, p. Mo3B.2, Optica Publishing Group, 2022.

[11] Z. Zhong, M. Ghobadi, M. Balandat, S. Katti, A. Kazerouni, J. Leach, M. McKillop, and Y. Zhang, "Bow: First real-world demonstration of a bayesian optimization system for wavelength reconfiguration," in *Optical Fiber Communication Conference (OFC) 2021*, p. F3B.1, Optica Publishing Group, 2021.

[12] A. Ferrari, V. V. Garbhapu, D. L. Gac, I. F. de Jauregui Ruiz, G. Charlet, and Y. Pointurier, "Demonstration of AI-Light: an Automation Framework to Optimize the Channel Powers Leveraging a Digital Twin," in *Optical Fiber Communication Conference (OFC) 2022*, p. M3Z.14, Optica Publishing Group, 2022.

[13] X. Yang, A. Ferrari, D. L. Gac, G. Charlet, M. Tornatore, and Y. Pointurier, "Experimental impact of power re-optimization in a mesh network," *Journal of Optical Communications and Networking*, vol. 15, no. 7, pp. C20–C28, 2023.

[14] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-Model of Fiber Non-Linear Propagation and its Applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, 2014.

[15] A. Ferrari, *Physical Layer Aware Open Optical Networking*. PhD thesis, Politecnico di Torino, Turin, Italy, 2021.

[16] H. Buglia, R. I. Killey, and P. Bayvel, "Ultra-wideband modeling of optical fibre nonlinearity in hybrid-amplified links," *Journal of Lightwave Technology*, vol. 42, no. 24, pp. 8664–8677, 2024.

[17] K. Nikolaou, D. Uzunidis, F. Arpanaei, J. M. Rivas-Moscoso, D. Larrabeiti, and I. Tomkos, "Maximizing the transport capacity of optical multi-band wdm systems through power optimization," in *2024 International Conference on Optical Network Design and Modeling (ONDM)*, 2024.

[18] G. Borraccini, S. Straullu, A. D'Amico, F. Aquilino, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Local and global optimization methods for optical line control based on quality of transmission," *Journal of Optical Communications and Networking*, vol. 16, no. 5, pp. B60–B70, 2024.

[19] Y. Zhang, X. Pang, Y. Song, Y. Wang, Y. Zhou, H. Zhu, L. Zhang, Y. Fan, Z. Guo, S. Huang, M. Zhang, and D. Wang, "Optical Power Control for GSNR Optimization Based on C+L-Band Digital Twin Systems," *Journal of Lightwave Technology*, vol. 42, no. 1, pp. 95–105, 2024.

[20] X. Liu, Y. Zhang, M. Cai, Y. Liu, L. Yi, W. Hu, and Q. Zhuge, "SMOF: Simultaneous Modeling and Optimization Framework for Raman Amplifiers in C+L-Band Optical Networks," *Journal of Lightwave Technology*, vol. 42, no. 9, pp. 3174–3183, 2024.

[21] J.-P. Elbers and C. Fürst, "Chapter 7 - Spectral Power Fluctuations in DWDM Networks Caused by Spectral-Hole Burning and Stimulated Raman Scattering," in *Optically Amplified WDM Networks* (J. Zyskind and A. Srivastava, eds.), pp. 201–219, Oxford: Academic Press, 2011.

[22] D. C. Kilper and C. A. White, "Chapter 8 - Amplifier Issues for Physical Layer Network Control," in *Optically Amplified WDM Networks* (J. Zyskind and A. Srivastava, eds.), pp. 221–251, Oxford: Academic Press, 2011.

[23] J. Junio, D. C. Kilper, and V. W. S. Chan, "Channel Power Excursions From Single-Step Channel Provisioning," *Journal of Optical Communications and Networking*, vol. 4, no. 9, pp. A1–A7, 2012.

[24] C. Sun, X. Yang, G. Charlet, P. A. Stavrou, and Y. Pointurier, "Digital twin-enabled multi-step strategies for autonomous power equalization in optical networks," *Journal of Optical Communications and Networking*, vol. 17, no. 7, pp. C41–C50, 2025.

[25] C. Sun, X. Yang, G. Charlet, P. A. Stavrou, and Y. Pointurier, "Digital Twin-Enabled Optical Network Channel Power Management by WSS and Booster Auto-Adjustment," in *Optical Fiber Communication Conference (OFC) 2025*, p. W2A.44, Optica Publishing Group, 2025.

[26] C. Sun, X. Yang, G. Charlet, P. A. Stavrou, and Y. Pointurier, "Digital Twin Enabled Automatic Power Adjustment with Multi-Step Lookahead Prediction," in *50th European Conference on Optical Communication (ECOC)*, p. M2E.1, 2024.

[27] C. Sun, X. Yang, G. Charlet, P. A. Stavrou, and Y. Pointurier, "Digital Twin-Enabled Optical Network Automation: Power Re-Optimization," in *Optical Fiber Communication Conference (OFC) 2024*, p. W4I.5, Optica Publishing Group, 2024.

[28] C. Sun, X. Yang, G. Charlet, P. A. Stavrou, and Y. Pointurier, "Digital Twin-Enabled Service Optimization Sequence of Actions for Power Equalization," in *2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)*, 2023.

[29] C. Sun, X. Yang, N. D. Cicco, R. Ayassi, V. V. Garbhapu, P. A. Stavrou, M. Tornatore, G. Charlet, and Y. Pointurier, "Experimental Demonstration of Local AI-Agents for Lifecycle Management and Control Automation of Optical Networks," *Journal of Optical Communications and Networking*, vol. 17, no. 8, pp. C82–C92, 2025.

[30] C. Sun, X. Yang, N. D. Cicco, R. Ayassi, V. V. Garbhapu, P. A. Stavrou, M. Tornatore, G. Charlet, and Y. Pointurier, "First Experimental Demonstration of Full Lifecycle Automation of Optical Network through Fine-Tuned LLM and Digital Twin," in *50th European Conference on Optical Communication (ECOC) 2024*, p. Th3B.6, 2024.

[31] C. Sun, R. Ayassi, X. Yang, G. Charlet, P. A. Stavrou, and Y. Pointurier, "Demonstration of LLM-based AI-Agent for Optical Network Management and Automation," in *ECOC 2024; 50th European Conference on Optical Communication*, pp. 248–248, 2024.

[32] H. Huang, R. Ayassi, D. Wang, Y. Zhao, C. Sun, M. Zuo, Y. Xu, J. Wang, Y. Pointurier, G. Charlet, D. Zhang, and H. Li, "Field Trial of LLM-Based Autonomous Network Management with AI-Agent in Real-Time 400G/800G

Elastic Optical Network," in *2025 European Conference on Optical Communications (ECOC)*, pp. 1–4, 2025.

[33] E. Liamzon, "e& UAE and Huawei Drive AI Infrastructure with New Optical Innovation Center," 2025.

[34] J. C. Maxwell, *A treatise on electricity and magnetism*, vol. 1. Clarendon press, 1873.

[35] D. J. Griffiths, *Introduction to electrodynamics*. Cambridge University Press, 2023.

[36] M. Planck, "Ueber das Gesetz der Energieverteilung im Normalspectrum," *Annalen der Physik*, vol. 309, no. 3, pp. 553–563, 1901.

[37] K. Kao and G. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," *Proceedings of the Institution of Electrical Engineers*, vol. 113, pp. 1151–1158, 1966.

[38] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[39] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*, pp. 195–211, Springer, 2000.

[40] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Soviet Physics-JETP*, vol. 38, no. 2, pp. 248–253, 1974.

[41] F. Poletti, "Nested antiresonant nodeless hollow core fiber," *Optics Express*, vol. 22, no. 20, pp. 23807–23828, 2014.

[42] M. Petrovich, E. Numkam Fokoua, Y. Chen, H. Sakr, A. I. Adamu, R. Hassan, D. Wu, R. Fatobene Ando, A. Papadimopoulos, S. R. Sandoghchi, G. Jasion, and F. Poletti, "Broadband optical fibre with an attenuation lower than 0.1 decibel per kilometre," *Nature Photonics*, 2025.

[43] J. Cho and P. J. Winzer, "Probabilistic Constellation Shaping for Optical Fiber Communications," *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1590–1607, 2019.

[44] A. K. Srivastava and Y. Sun, "Chapter 4 - Advances in Erbium-Doped Fiber Amplifiers," in *Optical Fiber Telecommunications IV-A (Fourth Edition)* (I. P.

Kaminow and T. Li, eds.), Optics and Photonics, pp. 174–212, Burlington: Academic Press, fourth edition ed., 2002.

[45] M. Premaratne and G. P. Agrawal, *Semiconductor optical amplifiers*, p. 143–172. Cambridge University Press, 2011.

[46] C. V. Raman and K. S. Krishnan, "A New Type of Secondary Radiation," *Nature*, vol. 121, pp. 501–502, 1928.

[47] P. Hansen and L. Eskildsen, "Remote Amplification in Repeaterless Transmission Systems," *Optical Fiber Technology*, vol. 3, no. 3, pp. 221–237, 1997.

[48] D. M. Baney, P. Gallion, and R. S. Tucker, "Theory and Measurement Techniques for the Noise Figure of Optical Amplifiers," *Optical Fiber Technology*, vol. 6, no. 2, pp. 122–154, 2000.

[49] H. A. Haus, "The Proper Definition of Noise Figure of Optical Amplifiers," in *Optical Amplifiers and Their Applications*, p. OMB7, Optica Publishing Group, 2000.

[50] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity Limits of Optical Fiber Networks," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, 2010.

[51] S. J. Savory, R. J. Vincent, and D. J. Ives, "Design considerations for low-margin elastic optical networks in the nonlinear regime [invited]," *Journal of Optical Communications and Networking*, vol. 11, no. 10, pp. C76–C85, 2019.

[52] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-Model of Fiber Non-Linear Propagation and its Applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, 2014.

[53] D. Semrau, R. I. Killey, and P. Bayvel, "A closed-form approximation of the gaussian noise model in the presence of inter-channel stimulated raman scattering," *Journal of Lightwave Technology*, vol. 37, no. 9, pp. 1924–1936, 2019.

[54] I. F. de Jauregui Ruiz, A. Ghazisaeidi, T. Zami, S. Louis, and B. Lavigne, "An accurate model for system performance analysis of optical fibre networks with in-line filtering," in *45th European Conference on Optical Communication (ECOC)*, 2019.

[55] K. Ishii, J. Kurumida, and S. Namiki, "Experimental Investigation of Gain Offset Behavior of Feedforward-Controlled WDM AGC EDFA Under Various Dynamic Wavelength Allocations," *IEEE Photonics Journal*, vol. 8, no. 1, pp. 1–13, 2016.

[56] D. Kilper, S. Chandrasekhar, and C. White, "Transient gain dynamics of cascaded erbium doped fiber amplifiers with re-configured channel loading," in *2006 Optical Fiber Communication Conference and the National Fiber Optic Engineers Conference*, 2006.

[57] B. Birand, H. Wang, K. Bergman, D. Kilper, T. Nandagopal, and G. Zussman, "Real-Time Power Control for Dynamic Optical Networks—Algorithms and Experimentation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 8, pp. 1615–1628, 2014.

[58] J. Tsai, Z. Wang, Y. Pan, D. C. Kilper, and L. Pavel, "Stability analysis in a ROADM-based multi-channel quasi-ring optical network," *Optical Fiber Technology*, vol. 21, pp. 40–50, 2015.

[59] D. Wang, Y. Song, Y. Zhang, X. Jiang, J. Dong, F. N. Khan, T. Sasai, S. Huang, A. P. T. Lau, M. Tornatore, and M. Zhang, "Digital Twin of Optical Networks: A Review of Recent Advances and Future Trends," *Journal of Lightwave Technology*, vol. 42, no. 12, pp. 4233–4259, 2024.

[60] Q. Zhuge, X. Liu, Y. Zhang, M. Cai, Y. Liu, Q. Qiu, X. Zhong, J. Wu, R. Gao, L. Yi, and W. Hu, "Building a digital twin for intelligent optical networks [Invited Tutorial]," *Journal of Optical Communications and Networking*, vol. 15, no. 8, pp. C242–C262, 2023.

[61] R. Vilalta, L. Gifre, R. Casellas, R. Muñoz, R. Martínez, A. Mozo, A. Pastor, D. López, and J. P. Fernández-Palacios, "Applying Digital Twins to Optical Networks with Cloud-Native SDN Controllers," *IEEE Communications Magazine*, vol. 61, no. 12, pp. 128–134, 2023.

[62] M. S. Faruk and S. J. Savory, "Measurement Informed Models and Digital Twins for Optical Fiber Communication Systems," *Journal of Lightwave Technology*, vol. 42, no. 3, pp. 1016–1030, 2024.

[63] P. Vizarreta, C. M. Machuca, and W. Kellerer, "Controller placement strategies for a resilient SDN control plane," in *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*, pp. 253–259, 2016.

[64] Y. Li and D. C. Kilper, "Optical Physical Layer SDN [Invited]," *Journal of Optical Communications and Networking*, vol. 10, no. 1, pp. A110–A121, 2018.

[65] K. Christodoulopoulos, C. Delezoide, N. Sambo, A. Kretsis, I. Sartzetakis, A. Sgambelluri, N. Argyris, G. Kanakis, P. Giardina, G. Bernini, D. Roccato, A. Percelsi, R. Morro, H. Avramopoulos, P. Castoldi, P. Layec, and S. Bigo, "Toward efficient, reliable, and autonomous optical networks: the ORCHESTRA solution [Invited]," *Journal of Optical Communications and Networking*, vol. 11, no. 9, pp. C10–C24, 2019.

[66] A. Mahajan, K. K. Christodoulopoulos, R. Martínez, R. Muñoz, and S. Spadaro, "Quality of transmission estimator retraining for dynamic optimization in optical networks," *Journal of Optical Communications and Networking*, vol. 13, no. 4, pp. B45–B59, 2021.

[67] H. Zhang, F. Gao, J. Cheng, L. Dou, S. Chen, B. Yan, Z. Sun, L. Wang, and C. Xie, "Demonstration of a Disaggregated ROADM Network with Automatic Channel Provisioning and Link Power Adjustment," in *2021 European Conference on Optical Communication (ECOC)*, 2021.

[68] T. Inoue, E. Shibano, H. Taga, and K. Goto, "In-Service Upgrade Method for DWDM Submarine Cable System Using ASE Dummy Lights," in *Optical Fiber Communication Conference*, p. MF67, Optica Publishing Group, 2004.

[69] D. J. Elson, G. Saavedra, K. Shi, D. Semrau, L. Galdino, R. Killey, B. C. Thomsen, and P. Bayvel, "Investigation of bandwidth loading in optical fibre transmission using amplified spontaneous emission noise," *Optics Express*, vol. 25, no. 16, pp. 19529–19537, 2017.

[70] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[71] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019, 2021.

[72] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[73] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[74] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672–2680, 2014.

[75] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998–6008, 2017.

[76] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[77] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *International Conference on Learning Representations*, 2014.

[78] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.

[79] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958.

[80] D. Rafique and L. Velasco, "Machine learning for network automation: overview, architecture, and applications [Invited Tutorial]," *Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D126–D143, 2018.

[81] D. Côté, "Using machine learning in communication networks [Invited]," *Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D100–D109, 2018.

[82] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An Overview on Application of Machine Learning Techniques in Optical Networks," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1383–1408, 2019.

[83] Z. Wang, D. C. Kilper, and T. Chen, "Open EDFA gain spectrum dataset and its applications in data-driven EDFA gain modeling," *Journal of Optical Communications and Networking*, vol. 15, no. 9, pp. 588–599, 2023.

[84] X. Jiang, J. Dong, Y. Song, J. Li, M. Zhang, and D. Wang, "Physics-Informed Machine Learning for EDFA: Parameter Identification and Gain Estimation," *Journal of Lightwave Technology*, vol. 43, no. 11, pp. 5040–5054, 2025.

[85] C. Harvey and S. J. Savory, "Optical power spectrum prediction using cascaded learning with uncertainty propagating noisy input gaussian processes," *Journal of Lightwave Technology*, vol. 43, no. 16, pp. 7639–7649, 2025.

[86] F. Musumeci, C. Rottondi, G. Corani, S. Shahkarami, F. Cugini, and M. Tornatore, "A Tutorial on Machine Learning for Failure Management in Optical Networks," *Journal of Lightwave Technology*, vol. 37, no. 16, pp. 4125–4139, 2019.

[87] D. Wang, C. Zhang, W. Chen, H. Yang, M. Zhang, and A. P. T. Lau, "A review of machine learning-based failure anagement in optical networks," *Science China Information Sciences*, vol. 65, no. 11, p. 211302, 2022.

[88] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, "Machine-learning-based anomaly detection in optical fiber monitoring," *Journal of Optical Communications and Networking*, vol. 14, no. 5, pp. 365–375, 2022.

[89] S. K. Singh and A. Jukan, "Machine-Learning-Based Prediction for Resource (Re)allocation in Optical Data Center Networks," *Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D12–D28, 2018.

[90] P. Freire, E. Manuylovich, J. E. Prilepsky, and S. K. Turitsyn, "Artificial neural networks for photonic applications—from algorithms to implementation: tutorial," *Adv. Opt. Photon.*, vol. 15, no. 3, pp. 739–834, 2023.

[91] Z. Shabka and G. Zervas, "Network-aware compute and memory allocation in optically composable data centers with deep reinforcement learning and graph neural networks," *Journal of Optical Communications and Networking*, vol. 15, no. 2, pp. 133–143, 2023.

[92] S. K. Patri, A. Paz, M. Wenning, J. Müller, and C. Mas-Machuca, "Long-term upgrade strategies in multiband and multifiber optical transport networks," *Journal of Optical Communications and Networking*, vol. 16, no. 5, pp. 602–612, 2024.

[93] Y. Wang, C. Zhang, J. Li, Y. Pang, L. Zhang, M. Zhang, and D. Wang, "AlarmGPT: an intelligent alarm analyzer for optical networks using a generative pre-trained transformer," *Journal of Optical Communications and Networking*, vol. 16, no. 6, pp. 681–694, 2024.

[94] D. Wang, Y. Wang, X. Jiang, Y. Zhang, Y. Pang, and M. Zhang, "When Large Language Models Meet Optical Networks: Paving the Way for Automation," *Electronics*, vol. 13, no. 13, 2024.

[95] X. Jiang, M. Zhang, Y. Song, Y. Zhang, Y. Wang, C. Ju, and D. Wang, "OptiComm-GPT: a GPT-based versatile research assistant for optical fiber communication systems," *Optics Express*, vol. 32, no. 12, pp. 20776–20796, 2024.

[96] Y. Pang, M. Zhang, Y. Liu, X. Li, Y. Wang, Y. Huan, Z. Liu, J. Li, and D. Wang, "Large language model-based optical network log analysis using LLaMA2 with instruction tuning," *Journal of Optical Communications and Networking*, vol. 16, no. 11, pp. 1116–1132, 2024.

[97] D. Adanza, C. Natalino, L. Gifre, R. Muñoz, P. Alemany, P. Monti, and R. Vilalta, "IntentLLM: An AI Chatbot to Create, Find, and Explain Slice Intents in TeraFlowSDN," in *2024 IEEE 10th International Conference on Network Softwarization (NetSoft)*, pp. 307–309, 2024.

[98] A. Zhou, Y. Song, Y. Zhang, M. Zhang, and D. Wang, "Large Language Model-Driven AI Agent in SDN Controller Towards Intent-Based Management of Optical Networks," in *50th European Conference on Optical Communication (ECOC) 2024*, 2024.

[99] Y. Song, Y. Zhang, A. Zhou, Y. Shi, S. Shen, X. Tang, J. Li, M. Zhang, and D. Wang, "Synergistic Interplay of Large Language Model and Digital Twin for Autonomous Optical Networks: Field Demonstrations," *IEEE Communications Magazine*, 2024.

[100] N. D. Cicco, M. Ibrahimi, S. Troia, F. Musumeci, and M. Tornatore, "Open Implementation of a Large Language Model Pipeline for Automated Configuration of Software-Defined Optical Networks," in *50th European Conference on Optical Communication (ECOC) 2024*, 2024.

[101] V. Karunakaran, C. Natalino, B. Shariati, P. Lechowicz, J. K. Fischer, A. Autenrieth, P. Monti, and T. Bauschert, "TAPI-based Telemetry Streaming in Multi-domain Optical Transport Network," in *Optical Fiber Communication Conference (OFC) 2024*, p. M3Z.9, Optica Publishing Group, 2024.

[102] E. Etezadi, C. Natalino, V. Karunakaran, R. Diaz, A. Lindgren, S. Melin, A. Autenrieth, L. Wosinska, P. Monti, and M. Furdek, "Demonstration of

DRL-based intelligent spectrum management over a T-API-enabled optical network digital twin," in *49th European Conference on Optical Communications (ECOC 2023)*, vol. 2023, pp. 1730–1733, 2023.

[103] M. Devigili, M. Ruiz, N. Costa, C. Castro, A. Napoli, J. Pedro, and L. Velasco, "Applications of the OCATA time domain digital twin: from QoT estimation to failure management," *Journal of Optical Communications and Networking*, vol. 16, no. 2, pp. 221–232, 2024.

[104] OpenAI, "ChatGPT: Large Language Model." https://openai.com, 2024.

[105] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, J.-P. Guillou, S. Schmidt, and M. Auli, "Llama: Open and efficient foundation language models," 2023. arXiv:2302.13971.

[106] Mistral AI. https://mistral.ai, 2024.

[107] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-Rank Adaptation of Large Language Models," 2021. arXiv:2106.09685.

[108] Mistral AI, "Mistral-7b-instruct-v0.3." https://huggingface.co/mistralai/Mistral-7B-Instruct-v0.3.

[109] Mistral AI, "Mixtral-8x7b-instruct-v0.1." https://huggingface.co/mistralai/Mistral-8x7B-Instruct-v0.1.

[110] Hugging Face, "Hugging Face." https://huggingface.co.

[111] LangChain, "Langchain." https://github.com/langchain-ai/langchain.

[112] Chroma, "Chroma." https://github.com/chroma-core/chroma.

[113] M. Brysbaert, "How many words do we read per minute? A review and meta-analysis of reading rate," *Journal of Memory and Language*, vol. 109, p. 104047, 2019.