A Game Theoretical Approach to Constrained OSNR Optimization Problems in Optical Networks

by

Yan Pan

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Electrical and Computer Engineering
University of Toronto

Copyright © 2009 by Yan Pan
Abstract

A Game Theoretical Approach to Constrained OSNR Optimization Problems in Optical Networks

Yan Pan
Doctor of Philosophy
Graduate Department of Electrical and Computer Engineering
University of Toronto
2009

Optical signal-to-noise ratio (OSNR) is considered as the dominant performance parameter at the physical layer in optical networks. This thesis is interested in control and optimization of channel OSNR by using optimization and game-theoretic approaches, incorporating two physical constraints: the link capacity constraint and the channel OSNR target.

To start, we study OSNR optimization problems with link capacity constraints in single point-to-point fiber links via two approaches. We first present a framework of a Nash game between channels towards optimizing individual channel OSNR. The link capacity constraint is imposed as a penalty term to each cost function. The selfish behavior in a Nash game degrades the system performance and leads to the inefficiency of Nash equilibria. From the system point of view, we formulate a system optimization problem with the objectives of achieving an OSNR target for each channel while satisfying the link capacity constraint. As an alternative to study the efficiency of Nash equilibria, we use the system framework to investigate the effects of parameters in cost functions in the game-theoretic framework.

Then extensions to multi-link and mesh topologies are carried out. We propose a partition approach by using the flexibility of channel power adjustment at optical switches. The multi-link structure is partitioned into stages with each stage being a single sink. By
fully using the flexibility, a more natural partition approach is applied to mesh topologies where each stage is a single link. The closed loop in mesh topologies can be unfolded by selecting a starting link. Thus instead of maximization of channel OSNR from end to end, we consider minimization of channel OSNR degradation between stages. We formulate a partitioned Nash game which is composed of ladder-nested stage Nash games.

Distributed algorithms towards the computation of a Nash equilibrium solution are developed for all different game frameworks. Simulations and experimental implementations provide results to validate the applicability of theoretical results.
Acknowledgements

To my supervisor, Professor Lacra Pavel, I would like to express my most heartfelt gratitude. For me, she is a devoted teacher, a caring mentor and a role model for my academic life. I am grateful for the resources and opportunities she provided, and for her guidance, support and patience throughout my pursuit of this degree.

I would like to thank my advisory committee members, Professor Bruce A. Francis and Professor Shahrokh Valaee for their stimulating discussions and insightful comments on my research work. I would like to thank Professor Raymond H. Kwong and Professor Baochun Li for serving on my committee. Also I would like to thank Professor Ravi R. Mazumdar of the Department of Electrical and Computer Engineering at the University of Waterloo, for serving as an external examiner.

In addition, I would like to thank Tansu Alpcan in Deutsche Telekom Laboratories for helpful and stimulating discussions regarding the work in Chapter 6.

Special thanks to my colleagues in the Systems Control group. I am fortunate to be in such a dynamic group full of intelligent and pleasant people.

To my mom, for her care, love and support I shall be eternally thankful.
Contents

Abstract iv

Acknowledgements iv

List of Tables ix

List of Figures xiii

1 Introduction 1
  1.1 Scope of the Research ............................................. 3
  1.2 Related Work ..................................................... 7
  1.3 Main Contributions ............................................... 10
  1.4 Thesis Outline .................................................... 12

2 Background 13
  2.1 Introduction to Optimization .................................... 14
  2.2 Noncooperative (Nash) Game ................................... 18
     2.2.1 Nash Game with Uncoupled Constraints ..................... 18
     2.2.2 Nash Game with Coupled Constraints ....................... 26
  2.3 Optical Network Setup .......................................... 35
     2.3.1 OSNR Model ..................................................... 38
     2.3.2 Link Capacity Constraint .................................... 42
2.4 Summary ................................................................. 44

3 Games in Point-to-Point Topologies 45

3.1 Introduction .......................................................... 45
3.2 Game Formulation .................................................. 46
3.3 NE Solution .......................................................... 50
  3.3.1 Existence and Uniqueness .................................. 51
  3.3.2 Pricing Strategy ................................................ 55
  3.3.3 Properties of the NE Solution ............................. 56
3.4 Algorithms and Convergence Analysis ....................... 58
  3.4.1 Parallel Update Algorithm (PUA) ......................... 58
  3.4.2 Gradient Algorithm (GA) .................................. 64
3.5 Simulations .......................................................... 70
3.6 Summary ............................................................ 74

4 Games in Multi-link Topologies 75

4.1 Introduction ......................................................... 76
4.2 Games in Single-sink Multi-link Topologies ............... 77
  4.2.1 Convexity Analysis .......................................... 78
  4.2.2 Game Formulation ........................................... 80
  4.2.3 Decomposition ................................................ 81
4.3 Games in General Multi-link Topologies ................. 85
  4.3.1 Partitioned Game Formulation ............................ 88
  4.3.2 Iterative Hierarchical Algorithm ....................... 91
4.4 Simulations ........................................................ 102
4.5 Summary .......................................................... 107

5 Games in Mesh Topologies 108

5.1 Introduction ......................................................... 108
5.2 Network Model ......................................................... 110
5.3 Partitioned Game Formulation ................................. 114
5.4 Hierarchical Decomposition and Algorithm .................. 116
  5.4.1 Hierarchical Decomposition ................................. 117
  5.4.2 Hierarchical Algorithm ..................................... 119
5.5 Simulations ......................................................... 125
  5.5.1 Multi-link Topologies ....................................... 126
  5.5.2 Quasi-ring Topologies ..................................... 131
  5.5.3 Mesh Topologies ............................................. 133
5.6 Summary ......................................................... 138

6 Optimization with OSNR Target Constraints ...................... 143
  6.1 Introduction ..................................................... 144
  6.2 Problem Formulation ............................................ 145
    6.2.1 System Problem ............................................ 145
    6.2.2 Maximum OSNR Target ..................................... 150
  6.3 Distributed Algorithm .......................................... 151
    6.3.1 Relaxed System Problem ................................ 151
    6.3.2 Primal Algorithm ......................................... 153
  6.4 Numerical Study ................................................. 155
    6.4.1 Simulation Results of System Optimization Problem .... 156
    6.4.2 Comparison of Cost Functions ............................ 157
    6.4.3 Parameter Effects in the Nash Game ....................... 160
  6.5 Summary ....................................................... 167

7 Experiments ........................................................ 169
  7.1 Overview ....................................................... 169
  7.2 Implementations and Results ................................ 171
List of Tables

6.1 System cost values with different $\alpha_i$ ............................. 162
6.2 Parameters: $\beta_i$ ....................................................... 163
6.3 System cost values with different $\beta_i$ ................................. 164
# List of Figures

1.1 Block diagram of an all-optical WDM network system . . . . . . . . . . . 2
1.2 Network control block . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

2.1 Reaction curves, $r_1(u_2)$ and $r_2(u_1)$. . . . . . . . . . . . . . . . . . . . . . . 22
2.2 Constraints: Rectangular and triangular . . . . . . . . . . . . . . . . . . 27
2.3 Reaction curves in coupled Nash game: $r_1(u_2)$ and $r_2(u_1)$ . . . . . 29
2.4 A wavelength-routed optical network configuration: Mesh, link and span 35
2.5 A point-to-point WDM fiber link . . . . . . . . . . . . . . . . . . . . . . 38
2.6 Three-link topology with three channels . . . . . . . . . . . . . . . . . . 43

3.1 $U_i$ vs $OSNR_i$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
3.2 Illustration of $\tilde{u}_i$ and $\tilde{u}_i'$ . . . . . . . . . . . . . . . . . . . . . . . 58
3.3 Iteration for channel $i$, with $u_i(0) = 0$ . . . . . . . . . . . . . . . . . . 60
3.4 Distributed PUA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
3.5 PUA and GA: channel power . . . . . . . . . . . . . . . . . . . . . . . . . 71
3.6 PUA and GA: total power . . . . . . . . . . . . . . . . . . . . . . . . . . 71
3.7 PUA and GA: channel OSNR . . . . . . . . . . . . . . . . . . . . . . . . . 72
3.8 GA: channel power . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 72
3.9 GA: total power . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 73
3.10 GA: channel OSNR . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 73

4.1 A single-sink multi-link topology . . . . . . . . . . . . . . . . . . . . . . . 78
4.2 Example: Three-link topology with three channels .................................. 87
4.3 Channel algorithm (4.31) and link algorithm (4.41) applied on stage $k$ .... 100
4.4 Flow of the channel and link algorithm .................................................. 101
4.5 Evolution of input power on stage 1 ....................................................... 103
4.6 Evolution of OSNR on stage 1 ................................................................. 103
4.7 Evolution of total power and price on link 1, stage 1 ............................... 104
4.8 Evolution of total power and price on link 2, stage 1 ............................... 105
4.9 Evolution of input power on stage 2 ....................................................... 105
4.10 Evolution of OSNR on stage 2 ............................................................... 106
4.11 Evolution of total power and price on stage 2 ....................................... 106

5.1 Quasi-ring topology ............................................................................ 109
5.2 Link $l$ ................................................................................................. 111
5.3 Illustration for case 1 ........................................................................... 123
5.4 Illustration for case 2 ........................................................................... 125
5.5 Three-player game on a three-link topology ........................................... 127
5.6 Multi-link case: Channel input power on link 1 ..................................... 128
5.7 Multi-link case: Total power and link price on link 1 .............................. 128
5.8 Multi-link case: Channel input power on link 2 ..................................... 129
5.9 Multi-link case: Total power and link price on link 2 .............................. 129
5.10 Multi-link case: Channel input power on link 3 .................................. 130
5.11 Multi-link case: Total power and link price on link 3 ............................ 130
5.12 Unfolded quasi-ring topology with the starting link, link 3 ................. 132
5.13 Quasi-ring case: Channel input power on link 3 .................................. 133
5.14 Quasi-ring case: Total power and link price on link 3 ........................... 134
5.15 Quasi-ring case: Channel input power on link 1 .................................. 134
5.16 Quasi-ring case: Total power and link price on link 1 ........................... 135
5.17 Quasi-ring case: Channel input power on link 2 .................................. 135

xi
7.4  Evolution of total power: m=2 ............................................. 174
7.5  Snapshot on OSA: Results of PUA ........................................ 174
7.6  Evolution of input power with GA, m=5 ................................. 175
7.7  Evolution of total power with GA, m=5 ................................. 176
7.8  Snapshot on OSA: Results of GA, m=5 ................................. 177
7.9  Snapshot on OSA: Results of GA, m=5 ................................. 178
7.10 ONTS Configuration Setup: Multiple links ......................... 178
7.11 Evolution of input power on stage 1 .................................... 179
7.12 Evolution of total power and price on stage 1 ....................... 180
7.13 Evolution of input power on stage 2 .................................... 180
7.14 Evolution of total power and price on stage 2 ....................... 181
7.15 Snapshot: Partial evolution of OSNR and attenuation value on stage 2 . 181

A.1 Example: upper-semi-continuous ........................................... 192
Chapter 1

Introduction

Optical networks are high-capacity communication networks where optical fiber is used as the transmission medium. Wavelength division multiplexing (WDM) technology exploits high bandwidth by multiplexing multiple signals with different wavelengths into a single optical fiber. In first-generation optical networks, electronic switches were used and a transmitted signal might go through electrical-to-optical (E/O) and optical-to-electrical (O/E) conversion many times before reaching its destination. Optical networks are called all-optical if a transmitted signal remains in the optical form from its source (transmitter, Tx) to its destination (receiver, Rx) without undergoing any O/E conversions. Fig. 1.1 shows a block diagram of an all-optical WDM network. The network medium may be a single point-to-point WDM fiber link, or a network of optical switches and WDM fiber links. The transmitter block consists of one or more transmitters. Each transmitter consists of a laser and a laser modulator to carry out the E/O conversion, one for each wavelength. If multiple transmitters are used, then the signals at the different wavelengths are combined onto a single fiber. The receiver block may consist of a demultiplexer and a set of photodetectors (receivers). Each photodector receives a signal at a wavelength and carries out the O/E conversion. As the optical signal propagates through the fiber, it is attenuated; that is, its power decreased. Optical amplifiers (OAs)
are deployed along fiber links throughout the network to maintain the transmitted signal powers.

![Block diagram of an all-optical WDM network system](image)

Figure 1.1: Block diagram of an all-optical WDM network system

In recent years we have witnessed the evolution of research beyond point-to-point WDM fiber links to \textit{switched} (also referred to as \textit{wavelength-routed}) all-optical networks. Such a network includes a number of optical switches connected by fiber links to form a \textit{physical} topology (physical layer). Different optical switches include optical cross-connects (OXC\textsc{s}) and optical add/drop multiplexers (OAD\textsc{m}s). An OXC deals with multiple wavelengths and selectively drops some of these wavelengths locally or adds selected wavelengths while letting others pass through. An OADM provides a similar function but at a much smaller size. A \textit{lightpath} (also referred to as \textit{channel}) is established for each transmitted signal at a wavelength from its associated transmitter to the corresponding receiver. A lightpath is routed and switched by optical switches over a wavelength on each intermediate fiber link. The collection of lightpaths is called the \textit{virtual} topology (optical layer). The optical switch may be equipped with a wavelength converter. Then some channels may be converted from one wavelength to another as well
along their route by *wavelength conversion*. Optical switches can be used to amplify or attenuate channel signal powers individually [2].

The advancement of optical technologies makes it possible to deploy *reconfigurable* switched all-optical networks. In such a network, the network reconfiguration with arbitrary virtual topologies can be quickly performed under software control. This leads to an emerging and promising research goal:

- Realize reconfiguration with arbitrary virtual topologies;
- Keep network stability;
- Maintain optical channel transmission performance.

More precisely, transmitted signals degrade in quality due to physical layer impairments. The signal on the new established channel affects the quality of each existing signal on shared fiber links. This thesis is interested in the control of channel transmission performance incorporating physical layer impairments after reconfiguration by using optimization and game-theoretic approaches. For simplicity, we use “optical networks” to generally refer “reconfigurable switched all-optical networks”. Without causing any confusing, we also use “channel (input) power” to refer “channel (input) signal power”.

The remainder of this chapter is organized as follows. We start with the scope of present research work in Section 1.1. Section 1.2 provides an extensive review of the literature on control and optimization in communication networks. The main contributions of this thesis follow immediately in the next section. Finally, Section 1.4 summarizes the organization of the thesis.

### 1.1 Scope of the Research

In recent years, researchers have realized that optical networks can have more complicated topologies than just point-to-point WDM fiber links. In such networks, the length
Chapter 1. Introduction

of fiber links and the number of OAs deployed along a link are changing. Moreover this network provides switching and routing functions by using optical switches. Thus different channels may travel on different fiber links and arbitrary virtual topologies can be formed. In static optical networks, lightpaths are statically provisioned up front. The operational expense of reconfiguring the network is quite high and such reconfiguration is time-consuming in the sense that any change in the virtual topology requires manual reconfiguration and additional equipments.

Today it is of great interest and demand to deploy reconfigurable optical networks in which quick reconfiguration is realized under software control and at the same time, the network stability and channel transmission performance are maintained. Evolution of emerging reconfigurable network elements (such as reconfigurable OADMs, OXCs and tunable lasers) make it possible to realize such networks. The published work on network performance by using these reconfigurable elements can be found in [71, 79, 84, 92] and references therein. Instead of looking from the viewpoint of the wavelength routing or wavelength assignment [53, 67, 91] in the context of reconfigurable optical networks, our research objective in this thesis is to control and optimize channel transmission performance (also referred to as quality of service, QoS) at the physical layer.

While a channel traverses a number of optical elements (e.g., OAs and OXCs) through the network, the channel signal encounters crosstalk at OXCs and amplified spontaneous emission (ASE) noise at OAs, thus it may degrade in quality due to these impairments. Channel transmission performance at the physical layer is characterized through the bit-error rate (BER), which is defined as the average probability of incorrect bit identification. Thus performance degradation is associated with an increase in the BER. The BER in turn depends on optical signal-to-noise ratio (OSNR), dispersion and various nonlinear effects [2].

Indeed, optical network system performance is limited by dispersive and nonlinear effects. The effects of dispersion can be minimized by using a dispersion management
technique [2]. The impairment due to nonlinear effects is an important consideration in the design of an amplified fiber link. Effects of nonlinearity accumulate when OAs are used periodically to compensate signal power attenuation on fibers. The nonlinear effects are negligible if the fiber link is designed properly with optimized total input power of all signals. Such a condition is regarded as a link capacity constraint. The effects of nonlinearity can also be minimized by a map of OA gain distribution proposed by Mecozzi in [46]. Thus network system performance can be improved by dispersion management technique and optimization of the total input power. For further details on fiber nonlinearities, dispersion and their effects, readers are referred to [2, 11, 17, 18, 34, 48, 66, 79, 93].

Thus OSNR becomes the dominant performance parameter when nonlinear and dispersive impairments are limited. This is called the noise-limited regime versus the dispersion-limited regime. These two regimes can be treated separately. ASE noise generated and accumulated in chains of OAs degrades associated channel OSNR. Detailed OSNR computations are needed to determine whether the channel OSNR on a given optical device or route is acceptable [79].

The objective to control and optimize channel transmission performance is directly translated to how to control and optimize channel OSNR. Note that an increase of the input power of a channel leads to the improvement of its OSNR, but high total input power of all channels will cause the nonlinear effects to be dominant. This trade-off imposes a constraint on the total input power while optimizing channel OSNR. Therefore, the work of this thesis is concerned with how to solve the constrained OSNR optimization problem in optical networks with arbitrary virtual topologies.

Particularly in the constrained OSNR optimization problem, we are interested the development of iterative power control algorithms based on local network measurements (e.g., OSNR), such that if channel input powers adjust, the goal of optimal OSNR is achieved. It is also desirable for such algorithms to allow for distributed processing with
respect to channels, bypassing the need of signal central processing units. Fig. 1.2 depicts one such feedback system with a power control algorithm. We take channel signal power at the sources (channel power) as inputs to the optical network system and channel OSNR outputs and other network measurements as feedback signals.

![Network control block](image)

Figure 1.2: Network control block

It is worth mentioning that in real-time, the power control algorithm is implemented in an iterative feedback manner. Each channel OSNR is measured by the optical spectrum analyzer (OSA) at the optical switches and the destination. Feedback network measurements are transmitted via optical service channels (OSCs) in optical networks. Each source has a computation unit which performs an algorithmic calculation and determines the corresponding channel input power.

Two questions remain to be answered: How to formulate meaningful and tractable optimization problems for such networks, and how to develop iterative, distributed algorithms with proven convergence? The answers are inspired by the following literature review on control and optimization in both optical networks and general communication networks.
Chapter 1. Introduction

1.2 Related Work

In the literature, OSNR consideration is not a completely new research topic in optical networks, particularly in point-to-point WDM fiber links [19, 86]. A simple link OSNR model was first proposed by Forghieri et al. in [27] for single point-to-point WDM fiber links. Based on this, Pavel [62] developed an analytical network OSNR model for general network configurations, which is the basis of our research and is reviewed in Chapter 2. On the other hand, mathematically the similarity between the OSNR model and the signal-to-interference ratio (SIR) model in wireless networks [5] inspires the idea that a similar methodology can be applied to solve OSNR-related problems in optical networks. Questions remain: In what respects are optical networks different from wireless networks and can existent results be applied “directly”? The first step we take to answer above questions is to provide an extensive literature review in this context.

Although the study on OSNR equalization can be traced back to early 1990’s, the published literature to date on channel performance at the physical layer from the theoretical perspective of OSNR optimization is relatively sparse.

A traditional approach used a static budget of fiber and device impairments along a fiber link with sufficient tolerance margins added [68]. Therefore at least a desired OSNR is achieved for each channel. The process is repeated manually after reconfiguration. Mecozzi in [46] proposed an optimum map of the OA gain distribution. The map not only minimized the effects of the nonlinearity, but also gave rise to the minimum OSNR for a given average power and the minimum average power for a given OSNR.

Power control has been used to solve OSNR equalization problems since 1990’s. Just as the name implies, the primary objective of power control is to obtain a satisfactory channel OSNR by regulating channel input power. Chraplyvy et al. in [19] proposed an online algorithm to equalize the OSNR of all channels by adjusting the individual input powers. Experimental and simulation work was performed in [20]. An analysis of the performance limits associated with the end-to-end equalization was presented by Tonguz.
et al. in [82, 83]. The algorithm in [19] was simple but heuristic and no convergence results were given. The algorithm worked in a static and centralized way in the sense that each channel required OSNR of all other channels and network reconfiguration was not considered. Another static OSNR equalization was formulated in [27] by Forghieri et al. They proposed a simple OSNR model in the form of a system transmission matrix. The optimal input powers were found by solving an eigenvalue problem.

The referred OSNR optimization approaches were developed for single point-to-point WDM fiber links and were not appropriate for optical networks. In fact, dynamic power control that optimally adjusts network parameters (channel power) based on online network feedback (from Rx to Tx and various optical switches) can result in increased flexibility, scalability and capacity of optical networks. Two critical issues are arising: meaningful problem formulations for OSNR optimization and development of online distributed algorithms with provable convergence properties.

To start, an optical network is considered as a dynamic multiuser environment, in which the signal of a channel on the same fiber link is regarded as an interfering noise for all others, which leads to OSNR degradation of other channels. Power control is a key issue in the design of an interference-limited multiuser communication network system.

Noncooperative (Nash) game theory [12,51,69] has been revealed to be a powerful tool to solve power control problems in different multiuser environments. In a Nash game, each player pursues to achieve maximization of its own utility (or minimization of its own cost) in response to the actions of all other players. A solution of such a game is called a Nash equilibrium (NE) solution. Although the selfish behavior of players causes system performance loss in a Nash game [1,22,35,41,65], it has been shown in [35,73] that proper selection of network pricing mechanism can help preventing the degradation of network system performance.

Yu et al. in [88] developed a power allocation scheme based on a game-theoretic approach for digital subscriber line (DSL) systems. They modeled each link in the DSL
system as a player with a utility function defined as its own achievable rate. A more widely studied power control problem lies in wireless networks, [5, 10, 24, 28, 33, 47, 72, 87, 90]. In [47, 87], distributed power control schemes were introduced via centralized (system-based) approaches. As an alternative to traditional system-based optimization approaches, game-theoretic approaches have been used widely in recent years [5, 24, 33, 72]. For example, a game-theoretic framework for power control in wireless data networks was presented by Falomari et al. in [24]. In particular, pricing mechanism was used to improve efficiency. Ji et al. in [33] studied games with various types of utility functions for cellular radio systems.

Research work for wireless networks in this topic provides us with considerable insight. However, there exist several specific features combined together making power control problems in optical networks more challenging: amplified spans, multiple links, accumulation and self-generation of ASE noise, as well as crosstalk at optical switches.

Work on OSNR optimization has been extended to optical networks firstly by Pavel [61–64]. A system-based approach was used in [62] to minimize channel input power while each channel could maintain a desired individual OSNR. An online distributed algorithm was also proposed for general network configurations. In [61] a Nash game among channels was formulated towards OSNR maximization. This formulation provided an explicit closed-form of a unique NE solution of the Nash game and led to an iterative distributed algorithm for the computation of the NE solution.

However, the link capacity constraint was not considered in either [62] or [61]. Approaches to solve such an OSNR optimization problem with coupled constraints can be inspired by approaches for congestion control (flow optimization) problems in Internet. Congestion control problems in Internet were solved by utilizing Nash game frameworks with uncoupled utility functions [3, 7, 8, 13, 42, 70, 74, 85], or by applying system-based optimization approaches where duality and separability results were used [38, 44, 45]. It is worth mentioning that in OSNR optimization problems, channel utility functions are
typically coupled, which complicates analysis.

In [63] Pavel extended duality results in optimization theory [15] to a Nash game framework with coupled utility functions and coupled constraints, which will be widely used in the thesis. A hierarchical decomposition based on Lagrangian extension and duality was proposed to compute an NE solution. As an application of the theoretical results, [63] studied a Nash game with coupled constraints but only in point-to-point WDM fiber links.

Finally we summarize that the research on OSNR optimization is evolving beyond static OSNR equalization in point-to-point fiber links. While incorporating various physical constraints, research has been started on dynamic OSNR optimization with coupled constraints both in links and networks with many open problems.

1.3 Main Contributions

This thesis studies OSNR optimization problems in optical networks incorporating physical constraints, link capacity constraint and/or channel OSNR target. The main contributions are summarized as follows.

- In Chapter 3, we propose a framework of a Nash game among channels towards maximizing channel OSNR in single point-to-point WDM links. We impose the link capacity constraint as a regulation term to each channel cost function, which penalizes any violation of the constraint. We propose two iterative algorithms in a distributed way to find the NE solution.

- Beyond single point-to-point WDM links, we study OSNR optimization problems in optical networks with arbitrary topologies in Chapters 4 and 5. This extension is nontrivial due to the fact that link capacity constraints are not always convex. We investigate a single-sink multi-link structure where convexity of link capacity constraints is ensured. Based on this investigation and the flexibility of channel
power adjustment at optical switches, we propose partition approaches and apply them to solve OSNR optimization problems in optical networks.

- In Chapter 4, we partition the multi-link structure into stages. Each stage has a single-sink structure and channel powers on each stage are adjustable. Instead of maximization of channel OSNR from end to end, we consider minimization of channel OSNR degradation between stages. We formulate a novel partitioned Nash game which is composed of several ladder-nested stage Nash games. Each stage Nash game is formulated towards minimizing channel OSNR degradation. We use Lagrangian extension and decomposition results introduced in Section 2.2.2 to develop a distributed hierarchical iterative algorithm (link pricing algorithm and channel power control algorithm) towards computing an NE solution.

- An extension of Chapter 4 to mesh topologies is carried out in Chapter 5. We propose a more natural partition approach in which each stage is a single link. By selecting starting points (links) for each quasi-ring or ring structure, we unfold closed loops and formulate a partitioned Nash game composed of ladder-nested link Nash games towards minimizing channel OSNR degradation between links. The partitioned structure is regular and scalable and it benefits the development of iterative algorithms.

- In Chapter 6, we study OSNR optimization problems with link capacity constraints and channel OSNR targets from the system point of view, for point-to-point WDM links. A barrier function is used to relax the original system optimization problem and leads to the development of a primal algorithm in a distributed way. This method was originally proposed in [77] and was used in [6] to solve power control problems in multicell CDMA wireless networks. We adopt it for the first time to solve power control problems in optical networks. We propose an alternate approach to study the efficiency of the NE solution in the game theoretical framework.
proposed in Chapter 3. We use the system optimization framework to investigate
the effects of parameters in game cost functions and show that efficiency can be
improved by appropriate selection of parameters.

- Experimental implementations are carried out to validate the applicability of the-
  oretical results in Chapter 7.

1.4 Thesis Outline

The thesis is organized as follows. Chapter 2 provides an overview of mathematical
optimization and Nash game theory. In addition, some basic concepts in optical networks
are introduced, forming a basis for the problems addressed in the next four chapters. In
Chapter 3, the OSNR optimization problem with the link capacity constraint for point-to-
point WDM fiber links is studied via a game-theoretic approach. Extensions to two types
of network topologies, the multi-link topology and the mesh topology, are carried out in
Chapters 4 and 5, respectively. Channel OSNR target is considered in Chapter 6 for point-
to-point WDM fiber links. Chapter 6 studies the OSNR optimization problem with link
power capacity constraint and channel OSNR target from a system point of view. This
is followed by a comparison in simulation between two approaches proposed in Chapter 3
and Chapter 6 for point-to-point WDM fiber links. Chapter 7 provides experimental
results. We conclude the thesis in Chapter 8 by presenting a concise summary of primary
contributions and a brief discussion about future research work.
Chapter 2

Background

This chapter provides an overview of optimization theory, noncooperative (Nash) game theory and a background of optical network systems. Sections 2.1 and 2.2 present necessary concepts and results of optimization theory and game theory mainly adapted from [12, 15]. The concepts of OSNR model and link capacity constraint in optical networks are introduced in Section 2.3 as a preface to the remaining chapters.

Some preliminary notations are presented first and will be used in this and the following chapters.

Let \( a = [a_i] \) and \( b = [b_i] \) be \( n \)-dimensional vectors. We write \( a \geq b \) if all \( a_i \geq b_i \) and \( a > b \) if all \( a_i > b_i \). The reverse relations, \( \leq \) and \( < \), are defined similarly. The notations are applicable if \( b = 0 \). We use \( \text{diag}(a) \) to denote a diagonal matrix whose diagonal entries are elements of vector \( a = [a_i] \). A superscript \( T \) denotes the transpose operation.

For a twice continuously differentiable function \( f : \mathbb{R}^m \to \mathbb{R} \), its gradient at \( u \) is denoted by \( \nabla f(u) \), a column vector defined as

\[
\nabla f(u) = \begin{bmatrix}
\frac{\partial f(u)}{\partial u_1} \\
\vdots \\
\frac{\partial f(u)}{\partial u_m}
\end{bmatrix},
\]

(2.1)
while the Hessian of function \( f \), denoted by \( \nabla^2 f \), is given as

\[
\nabla^2 f(u) = \begin{bmatrix}
\frac{\partial^2 f(u)}{\partial u_1^2} & \frac{\partial^2 f(u)}{\partial u_1 \partial u_2} & \ldots & \frac{\partial^2 f(u)}{\partial u_1 \partial u_m} \\
\frac{\partial^2 f(u)}{\partial u_2 \partial u_1} & \frac{\partial^2 f(u)}{\partial u_2^2} & \ldots & \frac{\partial^2 f(u)}{\partial u_2 \partial u_m} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 f(u)}{\partial u_m \partial u_1} & \frac{\partial^2 f(u)}{\partial u_m \partial u_2} & \ldots & \frac{\partial^2 f(u)}{\partial u_m^2}
\end{bmatrix}
\]

(2.2)

For a vector-valued function \( f : \mathbb{R}^m \rightarrow \mathbb{R}^m \), with components \( f_1, \ldots, f_m \), which are twice continuously differentiable, its pseudo-gradient at \( u \) is defined as a column vector of the first-order partial derivatives of \( f_i(u) \) with respect to \( u_i \),

\[
\nabla f(u) := \begin{bmatrix}
\frac{\partial f_1(u)}{\partial u_1} \\
\vdots \\
\frac{\partial f_m(u)}{\partial u_m}
\end{bmatrix},
\]

(2.3)

which is the diagonal of the Jacobian matrix of \( f \). The Jacobian of \( \nabla f(u) \) with respect to \( u \) is denoted by \( \nabla^2 f(u) \),

\[
\nabla^2 f(u) := \begin{bmatrix}
\frac{\partial^2 f_1(u)}{\partial u_1^2} & \frac{\partial^2 f_1(u)}{\partial u_1 \partial u_2} & \ldots & \frac{\partial^2 f_1(u)}{\partial u_1 \partial u_m} \\
\frac{\partial^2 f_2(u)}{\partial u_2 \partial u_1} & \frac{\partial^2 f_2(u)}{\partial u_2^2} & \ldots & \frac{\partial^2 f_2(u)}{\partial u_2 \partial u_m} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 f_m(u)}{\partial u_m \partial u_1} & \frac{\partial^2 f_m(u)}{\partial u_m \partial u_2} & \ldots & \frac{\partial^2 f_m(u)}{\partial u_m^2}
\end{bmatrix}.
\]

(2.4)

For a two-argument continuously differentiable function \( f(u; x) \) with \( f : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \), its gradient is defined as a two-column matrix,

\[
\nabla f(u; x) = \begin{bmatrix}
\nabla_u f(u; x) \\
\nabla_x f(u; x)
\end{bmatrix}
\]

(2.5)

where \( \nabla_u f(u; x) \) and \( \nabla_x f(u; x) \) are the gradients defined in (2.1), with respect to the first argument \( u = [u_1, \ldots, u_m] \) and the second argument \( x = [x_1, \ldots, x_m] \), respectively.

### 2.1 Introduction to Optimization

This section provides an overview of optimization problems. In particular, the covered topics include unconstrained and constrained optimization problems, the Lagrange mul-
A mathematical optimization problem has the form
\[
\begin{align*}
\min & \quad J_0(u) \\
\text{subject to} & \quad u \in \Omega,
\end{align*}
\]  
(2.6)

where \( u = [u_1, \ldots, u_m]^T \) is the variable, \( \Omega \subseteq \mathbb{R}^m \) is the general constraint set. The problem (2.6) is called an unconstrained optimization problem if \( \Omega = \mathbb{R}^m \). The objective function \( J_0 : \Omega \rightarrow \mathbb{R} \) is the cost function.

A vector satisfying the constraints in (2.6) is called a feasible vector for this problem. A feasible vector \( u^{opt} \) is a local minimum of \( J_0(u) \) if there exists an \( \epsilon > 0 \) such that
\[
J_0(u^{opt}) \leq J_0(u), \quad \forall \ u \in \Omega \text{ with } \|u - u^{opt}\| < \epsilon
\]
where \( \| \cdot \| \) denotes the Euclidean norm. A feasible vector \( u^{opt} \) is called a global minimum of \( J_0(u) \), or a solution of the problem (2.6) if
\[
J_0(u^{opt}) \leq J_0(u), \quad \forall \ u \in \Omega
\]
and \( u^{opt} \) is called strict if the inequality above is strict for \( u \neq u^{opt} \).

The following is a standard result (combining Proposition 1.1.1 and Proposition 1.1.2 in [15]) regarding the solution of (2.6).

Proposition 2.1 ([15]) Let \( u^{opt} \) be a local minimum of \( J_0 : \Omega \rightarrow \mathbb{R} \). Assume that \( J_0 \) is continuously differentiable in an open set \( S \subseteq \Omega \) containing \( u^{opt} \). Then
\[
\nabla J_0(u^{opt}) = 0. \quad \text{(first-order necessary condition)}
\]  
(2.7)

If in addition \( J_0(u) \) is twice continuously differentiable within \( S \), then
\[
\nabla^2 J_0(u^{opt}) \text{ is positive semidefinite.} \quad \text{(second-order necessary condition)}
\]  
(2.8)

If in addition \( S \) is convex and \( J_0(u) \) is a convex function over \( S \), then (2.7) is a necessary and sufficient condition for \( u^{opt} \in S \) to be a global minimum of \( J_0(u) \) over \( S \).
Chapter 2. Background

The problem (2.6) is called a constrained optimization problem if $\Omega$ is a strict subset of $\mathbb{R}^m$, $\Omega \subset \mathbb{R}^m$. Throughout this section, it is assumed that $\Omega$ is convex. The following result follows directly from Proposition 2.1.1 in [15].

**Proposition 2.2** If $\Omega \subset \mathbb{R}^m$ is a convex and compact set and $J_0 : \Omega \to \mathbb{R}$ is a strictly convex function, then the problem (2.6) admits a unique global minimum.

For constrained optimization problems, a specific structure of $\Omega \subset \mathbb{R}$ constructed by inequalities is taken into account in this thesis. That is,

$$\min J_0(u)$$

subject to $g_r(u) \leq 0, \, r = 1, \ldots, R,$

(2.9)

where constraints $g_1(u), \ldots, g_R(u)$ are real-valued continuously differentiable functions defined from $\mathbb{R}^m$ to $\mathbb{R}$. In a compact form the general constraint can be written as

$$\Omega = \{u \in \mathbb{R}^m \mid g_r(u) \leq 0, \, r = 1, \ldots, R\}.$$

We denote this optimization problem (2.9) by $OPT(\Omega, J_0)$.

For any feasible vector $u$, the set of active inequality constraints is denoted by

$$A(u) = \{r \mid g_r(u) = 0, \, r = 1, \ldots, R\}.$$  

(2.10)

If $r \notin A(u)$, the constraint $g_r(u)$ is inactive at $u$. A feasible vector $u$ is said to be regular if the active inequality constraint gradients $\nabla g_r(u), \, r \in A(u)$, are linear independent.

The Lagrangian function $L : \mathbb{R}^{m+R} \to \mathbb{R}$ for problem (2.9) is defined as

$$L(u, \mu) = J_0(u) + \sum_{r=1}^{R} \mu_r g_r(u),$$

(2.11)

where $\mu_r, \, r = 1, \ldots, R$, are scalars. The following proposition (Proposition 3.3.1 in [15]) states necessary conditions for solving $OPT(\Omega, J_0)$ in terms of the Lagrangian function defined in (2.11).
Proposition 2.3 (Karush-Kuhn-Tucker (KKT) Necessary Condition) Let $u^{\text{opt}}$ be a local minimum of the problem (2.9). Assume that $u^{\text{opt}}$ is regular. Then there exists a unique vector $\mu^* = (\mu^*_1, \ldots, \mu^*_R)$, called a Lagrange multiplier vector, such that

\begin{align*}
\nabla_u L(u^{\text{opt}}, \mu^*) &= 0 \quad (2.12) \\
\mu^*_r &\geq 0, \forall r = 1, \ldots, R \quad (2.13) \\
\mu^*_r &= 0, \forall r \notin A(u^{\text{opt}}). \quad (2.14)
\end{align*}

Condition (2.14) in Proposition 2.3 is called the complementary slackness condition.

The next result (Proposition 3.3.4 in [15]) reviews general sufficient conditions, in terms of the Lagrangian function (2.11), for the following constrained optimization problem:

\begin{align*}
\min & \quad J_0(u) \\
\text{subject to} & \quad g_r(u) \leq 0, \ r = 1, \ldots, R, \\
& \quad u \in \mathcal{U},
\end{align*}

(2.15)

where $\mathcal{U} \subseteq \mathbb{R}^m$. The conditions are general since differentiability and convexity of $J_0$ and $g_r$, $r = 1, \ldots, R$, are not required. Meanwhile, $\mathcal{U}$ may be a strict subset of $\mathbb{R}^m$, $\mathcal{U} \subset \mathbb{R}^m$.

Proposition 2.4 (General Sufficiency Condition) Consider the problem (2.15). Let $u^{\text{opt}}$ be a feasible vector, together with a vector $\mu^* = [\mu^*_1, \ldots, \mu^*_R]^T$ that satisfies

\begin{align*}
\mu^*_r &\geq 0, \forall r = 1, \ldots, R \\
\mu^*_r &= 0, \forall r \notin A(u^{\text{opt}})
\end{align*}

and assume $u^{\text{opt}}$ minimizes the Lagrangian function $L(u, \mu^*)$ (2.11) over $u \in \mathcal{U}$, denoted as

\begin{align*}
u^{\text{opt}} \in \arg \min_{u \in \mathcal{U}} L(u, \mu^*).
\end{align*}

Then $u^{\text{opt}}$ is a global minimum of the problem (2.15).
Remark 2.1 If in addition, $J_0$ and $g_r$, $r = 1, \ldots, R$, are also convex and $\mathcal{U} = \mathbb{R}^m$, then
the Lagrangian function $L(u, \mu)$ is convex with respect to $u$. Therefore by Proposition 2.1, (2.16) is equivalent to the first-order necessary condition (2.12) in Proposition 2.3. Thus conditions (2.12)-(2.14) are also sufficient.

2.2 Noncooperative (Nash) Game

In the standard optimization problem, there is only one decision-maker who aims to minimize an objective function by choosing values of variables from a constrained set such that the system performance is optimized. It can happen that the optimal solution can favor one particular participator. Game theory involves multiple decision-makers and sees participators as competitors (players). A game consists of a set of players, an action set (also referred to as a set of strategies) available to those players and an individual objective function for each player. In a game, each player individually takes an optimal action which optimizes its own objective function and each player’s success in making decisions depends on the decisions of the others.

Noncooperative (Nash) game theory [12, 51] is a powerful tool to study strategic interactions among self-interested players. A game is called noncooperative if each player pursues its own interests which are partly conflicting with others’. It is assumed that each player acts independently without collaboration or communication with any of the others [51].

2.2.1 Nash Game with Uncoupled Constraints

In this section, we consider Nash games with uncoupled constraints. A Nash game consists of three components: A set of $\mathcal{M} = \{1, \ldots, m\}$ players, an action set $\Omega_i$ and a cost function $J_i$, i.e., the objective function for each player $i$, $i \in \mathcal{M}$. We take $\Omega_i = [m_i, M_i]$, where $m_i > 0$ and $M_i > m_i$ are scalars. Each player $i \in \mathcal{M}$ has an action
variable $u_i \in \Omega_i$. The **action space** $\Omega$ is defined by the Cartesian product of $\Omega_i$:

$$\Omega := \Omega_1 \times \cdots \times \Omega_m. \quad (2.17)$$

We also let $\Omega_{-i} := \Omega_1 \times \cdots \Omega_{i-1} \times \Omega_{i+1} \times \cdots \times \Omega_m$. It can be seen that $\Omega$ is compact, convex, and has a nonempty interior set. Moreover, the action space $\Omega$ is separable by construction since there is no coupling among $\Omega_i$. Any two players $i$ and $j$ can take their actions independently from separate action sets, $\Omega_i$ and $\Omega_j$, $j \neq i$, respectively. An **action vector** $u \in \Omega$ can be written as $u = [u_1, \ldots, u_m]^T$, or $u = (u_{-i}, u_i)$ with the vector $u_{-i} \in \Omega_{-i}$ obtained by deleting the $i^{th}$ element from $u$, i.e., $u_{-i} = [u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m]^T$.

An **individual cost function** $J_i : \Omega \to \mathbb{R}$ is assigned to each player $i$. Each player takes actions to minimize its own cost function $J_i(u_{-i}, u_i)$, regarding the actions of all other players. We denote such an $m$-player Nash game defined on separable action space $\Omega$ with cost functions $J_i$ by $GAME(\mathcal{M}, \Omega_i, J_i)$.

**Assumption 2.1** The cost function $J_i(u_{-i}, u_i)$ is twice continuously differentiable in all its arguments and strictly convex in $u_i$ for every $u_{-i} \in \Omega_{-i}$.

Given the actions of all other players, $u_{-i} \in \Omega_{-i}$, each player $i$ independently solves a constrained optimization problem, i.e.,

$$\begin{align*}
\min_{u_i} & \quad J_i(u_{-i}, u_i) \\
\text{subject to} & \quad u_i \in \Omega_i.
\end{align*} \quad (2.18)$$

The relevant concept in a Nash game is the **Nash equilibrium (NE)** solution. At an NE point, no player can benefit by altering its action unilaterally. Mathematically speaking, $u^* \in \Omega$ is an NE solution if $u^*_i$ is a solution to the optimization problem (2.18), given all other players taking equilibrium actions, $u^*_{-i}$.

**Definition 2.1 (NE)** A vector $u^* \in \Omega$ is called an **NE solution** of $GAME(\mathcal{M}, \Omega_i, J_i)$ if for all $i \in \mathcal{M}$ and for every $u^*_{-i} \in \Omega_{-i}$,

$$J_i(u^*_{-i}, u^*_i) \leq J_i(u^*_{-i}, u_i), \, \forall \ u_i \in \Omega_i. \quad (2.19)$$
If in addition \( u^* \) is not on the boundary of the action space \( \Omega \), it is called an inner NE solution.

Intuitively, for a given \( u_i \in \Omega_i \), player \( i \) takes an action \( \xi \) such that
\[
J_i(u_{-i}, \xi) \leq J_i(u_{-i}, u_i), \ \forall \ u_i \in \Omega_i.
\]
The reaction has a precise definition in Nash games.

**Definition 2.2 (Reaction function)** In \( GAME(\mathcal{M}, \Omega_i, J_i) \), \( \forall \ i \in \mathcal{M} \)
\[
R_i(u_{-i}) = \{ \xi \in \Omega_i : J_i(u_{-i}, \xi) \leq J_i(u_{-i}, u_i), \ \forall \ u_i \in \Omega_i \}\tag{2.20}
\]
is called the optimal reaction set of player \( i \). If \( R_i(u_{-i}) \) is a singleton\(^1\) for every given \( u_{-i} \in \Omega_{-i} \), then \( R_i(u_{-i}) \) is called the reaction function of player \( i \), and is specifically denoted by \( r_i : \Omega_{-i} \to \Omega_i \).

Some properties of reaction functions in \( GAME(\mathcal{M}, \Omega_i, J_i) \) are presented in the following proposition.

**Proposition 2.5** Consider \( GAME(\mathcal{M}, \Omega_i, J_i) \) with Assumption 2.1. For each \( i \in \mathcal{M} \), a unique continuous reaction function \( r_i : \Omega_{-i} \to \Omega_i \), exists.

**Proof:** Under Assumption 2.1, the cost function \( J_i(u_{-i}, u_i) \) is strictly convex in \( u_i \). By Proposition 2.2, there exists a unique minimum \( u_i \), for any given \( u_{-i} \), such that
\[
J_i(u_{-i}, u_i) < J_i(u_{-i}, \xi), \ \forall \ \xi \in \Omega_i, \ \xi \neq u_i.
\]
By Definition 2.2, this implies that a unique mapping, i.e., a reaction function \( r_i : \Omega_{-i} \to \Omega_i \) exists for each player \( i \). The continuity can be obtained directly from Berge’s Maximum Theorem (Theorem A.3).

Provided that reaction functions are well defined and a common intersection point of the reaction functions exists, the existence of an NE solution is established. This reasoning leads to the following proposition (adapted from Theorem 4.3 in [12]).

---

\(^1\)A singleton is a set with exactly one element.
**Proposition 2.6** Under Assumption 2.1, GAME($\mathcal{M}, \Omega_i, J_i$) admits an NE solution.

*Proof:* By strict convexity of the cost function and Proposition 2.5, for every given $u_{-i} \in \Omega_{-i}$, there exists a unique reaction function $r_i$ of player $i \in \mathcal{M}$ such that

$$u_i = r_i(u_{-i}) \in \Omega_i, \quad \forall \ i \in \mathcal{M}.$$  

The corresponding vector form is $u = r(u) \in \Omega$ with $r(u) = [r_1(u_{-1}), \ldots, r_m(u_{-m})]^T$. Furthermore by Proposition 2.5, $r_i$ is continuous, and so is $r(u)$. On the other hand, $r(u)$ maps a compact and convex set $\Omega$ into itself. From Brouwer’s fixed-point theorem (Theorem A.4), there exists a $u^*$ such that $u^* = r(u^*)$. By Definition 2.1, it readily follows that $u^*$ is an NE solution of GAME($\mathcal{M}, \Omega_i, J_i$).

The following result is immediate from Definition 2.1 and Proposition 2.1 regarding the innerness of an NE solution.

**Proposition 2.7** Under Assumption 2.1, if a vector $u^*$ is an inner NE solution of GAME($\mathcal{M}, \Omega_i, J_i$), then $u = u^*$ satisfies the following set of following necessary conditions:

$$\frac{\partial J_i}{\partial u_i}(u_{-i}, u_i) = 0, \quad \forall \ i \in \mathcal{M}. \quad (2.21)$$

Generally speaking, an NE solution is not an optimal solution of its associated optimization problem:

$$\min_{j \in \mathcal{M}} \sum_{j \in \mathcal{M}} J_j(u)$$

subject to $u \in \Omega$,

and vice versa. In fact it is well-known that in most cases Nash equilibria do not optimize overall system performance with the Prisoner’s Dilemma [52] being the best-known example. Let’s take a look at a two-player Nash game for illustration.
Example 2.1 In a two-player Nash game, the cost functions are defined by

\[
J_1(u_2, u_1) = 2u_1^2 - 2u_1 - u_1u_2
\]

\[
J_2(u_1, u_2) = u_2^2 - \frac{1}{2}u_2 - u_1u_2
\]

with \( u_i \in \Omega_i = [0, 8], \ i = 1, 2. \)

The constraint set \( \Omega = \Omega_1 \times \Omega_2 \) is convex and compact and has a nonempty interior set. From the cost functions, it follows that \( \frac{\partial^2 J_1}{\partial u_1^2} = 4 \) and \( \frac{\partial^2 J_2}{\partial u_2^2} = 2. \) Thus Assumption 2.1 is satisfied. From Proposition 2.5, the reaction functions \( r_i, i = 1, 2, \) exist and are continuous. By Definition 2.2, the reaction function \( r_1(u_2) \) can be obtained by optimizing \( J_1(u_2, u_1) \) with respect to \( u_1 \) for every given \( u_2. \) It follows that \( r_1(u_2) = \frac{1}{4}(u_2 + 2). \) Similarly, the reaction function \( r_2(u_1) \) is obtained as \( r_2(u_1) = \frac{1}{2}(u_1 + \frac{1}{2}). \) The reaction curves are shown in Fig. 2.1.

By Definition 2.1, an NE solution lies on both reaction curves. Therefore the intersection point of \( r_1(u_2) \) and \( r_2(u_1), \ (u_1^*, u_2^*) = (\frac{9}{11}, \frac{4}{7}), \) is an NE solution. The corresponding optimal cost values are \( J_1^* = -\frac{81}{98} \) and \( J_2^* = -\frac{16}{49}. \) Fig. 2.1 indicates that this game admits a unique NE solution.
Next consider the associated optimization problem

$$\min J_0(u_1, u_2)$$

subject to $0 \leq u_i \leq 8$, $i = 1, 2$,

where $J_0(u_1, u_2) = J_1(u_2, u_1) + J_2(u_1, u_2) = 2u_1^2 + u_2^2 - 2u_1 - \frac{1}{2}u_2 - 2u_1u_2$.

The constraint set $\Omega$ is still convex and compact. The Hessian of $J_0$ is positive definite for every $u \in \Omega$ with

$$\nabla^2 J_0 = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix}$$

such that the cost function $J_0$ is strictly convex over $\Omega$. From Proposition 2.2 the associated optimization problem admits a unique global minimum. By using Proposition 2.3, the optimal solution is $u^{opt} = (\frac{5}{4}, \frac{3}{2})$ and the optimal value is $J_0^{opt} = -\frac{13}{8}$.

It is obvious that $u^* \neq u^{opt}$. The following is a further discussion. In the game framework, given $u^{opt}$, the cost values are $J_1(u^{opt}) = -\frac{5}{4}$, which is less than $J_1(u^*)$ and $J_2(u^{opt}) = -\frac{3}{8}$ which is less than $J_2(u^*)$. It follows that with the presence of selfishness, the NE solution is not necessarily the optimal solution from a social point of view and the game settles down with less social revenue (less efficiency).

Next we provide an alternate definition of NE, in the sense of an augmented system-like cost function $\tilde{J} : \Omega \times \Omega \rightarrow \mathbb{R}$,

$$\tilde{J}(u; x) := \sum_{i=1}^{m} J_i(u_{-i}, x_i),$$

(2.22)

where $J_i(u_{-i}, u_i)$ is the individual cost function satisfying Assumption 2.1. We call this two-argument function $\tilde{J}$ the NG cost function.

NG cost functions can be used to solve Nash games with coupled constraints. In the following we provide an alternate proof of Proposition 2.6 by using the NG cost function. Also the necessary conditions for an inner NE solution of $GAME(M, \Omega_i, J_i)$ are reformulated with respect to the NG cost function.
Definition 2.3 (NE-NG Sense) A vector $u^* \in \Omega$ satisfying

$$\tilde{J}(u^*; u^*) \leq \tilde{J}(u^*; x), \quad \forall \ x \in \Omega \quad (2.23)$$

with $\tilde{J}$ defined in (2.22), is called an NE solution of $GAME(\mathcal{M}, \Omega_i, J_i)$.

Remark 2.2 The two definitions, Definitions 2.1 and 2.3 are equivalent. Note that (2.23) can be equivalently rewritten as for every given $u^* - i$,

$$\sum_{i=1}^{m} J_i(u^* - i, u^*_i) \leq \sum_{i=1}^{m} J_i(u^* - i, x_i), \quad \forall \ x \in \Omega.$$

It readily follows that $u^*$ satisfies Definition 2.3 if $u^*$ is an NE solution in the sense of Definition 2.1. Conversely, if $u^*$ satisfies Definition 2.3, it indeed constitutes an NE solution in the sense of Definition 2.1. This can be explained by assuming that, to the contrary, it does not. This implies that for some $i \in \mathcal{M}$, there exists a $\tilde{u}_i \neq u^*_i$ such that $J_i(u^* - i, \tilde{u}_i) < J_i(u^* - i, u^*_i)$. By adding $\sum_{j \in \mathcal{M}, j \neq i} J_j(u^* - j, u^*_j)$ to both sides, the following inequality holds,

$$\tilde{J}(u^*; \tilde{u}) < \tilde{J}(u^*; u^*),$$

where $\tilde{u} := (u^* - i, \tilde{u}_i)$ and $\tilde{u}_i \neq u^*_i$. This contradicts the hypothesis in (2.23).

Next we prove Proposition 2.6 by using the concept of the NG cost function. The proof essentially follows arguments similar to those used in Theorem 4.4 in [12], which have also been used by Pavel in [63]. We restate it here to give readers a deep impression of the two-argument NG cost function. The proof helps reformulate the necessary conditions (2.21) with respect to the NG cost function.

Proof of Proposition 2.6: From (2.22), the two-argument NG cost function $\tilde{J}(u; x)$ is separable in the second argument $x$ for every given $u$, i.e., each component cost function in $\tilde{J}(u; x)$ is decoupled in $x$ for every given $u$. Therefore by using (2.22), for every given $u$, the gradient of $\tilde{J}(u; x)$ with respect to $x$ is written as

$$\nabla_x \tilde{J}(u; x) := \begin{bmatrix}
\frac{\partial J_1}{\partial x_1}(u_{-1}, x_1) \\
\vdots \\
\frac{\partial J_m}{\partial x_m}(u_{-m}, x_m)
\end{bmatrix} \quad (2.24)$$
The Hessian of \( \tilde{J}(u;x) \) with respect to \( x \), \( \nabla_{xx}^2 \tilde{J}(u;x) \), is a diagonal matrix with the diagonal elements being \( \frac{\partial^2 J_i(u_{-i}, x_i)}{\partial x_i^2} \), \( i = 1, \ldots, m \). Under Assumption 2.1, it follows from strict convexity of \( J_i(u_{-i}, x_i) \) with respect to \( x_i \) (for every given \( u_{-i} \)) that \( \tilde{J}(u;x) \) is strict convex with respect to its second argument \( x \), for every given \( u \). Moreover, \( \tilde{J}(u;x) \) is continuous in its arguments. We define a reaction set:

\[
R(u) = \{ x \in \Omega | \tilde{J}(u;x) \leq \tilde{J}(u;w), \forall w \in \Omega \},
\]

or equivalently,

\[
R(u) = \arg \min_{x \in \Omega} \tilde{J}(u;x),
\]

where minimization on the right-hand side (RHS) is done with respect to the second argument \( x \) in \( \tilde{J}(u;x) \). Recall that \( \Omega \) is compact. By the continuity and convexity property of \( \tilde{J} \), it follows from Theorem A.3 (Berge’s Maximum Theorem) that \( R \) is an upper-semi-continuous mapping that maps each point \( u \) in \( \Omega \) into a compact and convex subset of \( \Omega \). Then by Kakutani Fixed-point Theorem (Theorem A.5), there exists a point \( u^* \) such that \( u^* \in R(u^*) \), i.e., \( u^* \) satisfies (2.23). By Definition 2.3, \( u^* \) is an NE solution of \( GAME(\mathcal{M}, \Omega_i, J_i) \).

Based on the proof, an NE solution \( u^* \) is a fixed-point solution that satisfies \( u^* \in R(u^*) \), i.e., \( u = u^* \) satisfies the implicit function:

\[
u = \arg \min_{x \in \Omega} \tilde{J}(u;x). \tag{2.25}\]

Moreover, if \( u^* \) is inner, \( u = u^* \) equivalently satisfies

\[
\nabla_x \tilde{J}(u;x) \bigg|_{x=u} = 0, \tag{2.26}\]

where the notation “\( |_{x=u} \)” denotes finding a fixed-point solution. By using (2.24), we get the component-wise form of (2.26):

\[
\frac{\partial J_i}{\partial x_i} (u_{-i}, x_i) \bigg|_{x_i=u_i} = 0, \forall i \in \mathcal{M}, \tag{2.27}\]
which are equivalent to the necessary conditions (2.21) with respect to $J_i$ presented in Proposition 2.7.

The procedure to find an inner NE solution with respect to $\tilde{J}$ is briefly summarized as follows. We first solve $\nabla_x \tilde{J}(u; x) = 0$ for every given $u$ (i.e., $x$ is the only variable), which gives $x$ as a function of $u$. Then we look for a fixed-point solution for these equations, which is denoted by $x = u$. Practically, solving the resulting set of $m$ equations, (2.27), leads to an inner NE solution. We give next an illustration of this procedure.

**Example 2.2** The two-player Nash game in Example 2.1 can be solved by using the following NG cost function,

$$\tilde{J}(u; x) = J_1(u_2, x_1) + J_2(u_1, x_2)$$

$$= (2x_1^2 - 2x_1 - x_1u_2) + (x_2^2 - \frac{1}{2}x_2 - u_1x_2) \quad (2.28)$$

Assume there exists an inner NE solution. Then given every $u$, in the first step we solve $\nabla_x \tilde{J}(u; x) = 0$, i.e.,

$$\frac{\partial J_1}{\partial x_1}(u_2, x_1) = 4x_1 - 2 - u_2 = 0$$

$$\frac{\partial J_2}{\partial x_2}(u_1, x_2) = 2x_2 - \frac{1}{2} - u_1 = 0.$$ 

This leads to $x$ as a function of $u$, i.e.,

$$x = \begin{bmatrix} 0 & \frac{1}{4} \\ \frac{1}{2} & 0 \end{bmatrix} u + \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$$

In the second step we solve for a fixed-point solution, by setting $x = u$ in the above, which leads to $(u_1^*, u_2^*) = (\frac{9}{14}, \frac{4}{7})$. 

\[\square\]

### 2.2.2 Nash Game with Coupled Constraints

In Nash games with uncoupled constraints, the action space is the Cartesian product of the individual action sets and players can affect only the cost functions of the other
players but not their feasible action sets. In Nash games with coupled constraints, each player’s action affects the feasible action sets of the other players. In Example 2.1, the action sets are $\Omega_i = [0, 8]$, $i = 1, 2$, such that the action space $\Omega$ is rectangular in $\mathbb{R}^2$. What if the action space is not rectangular but a subset of $\Omega$?

**Example 2.3** Consider the two-player Nash game in Example 2.1 with an action space, $\Omega := \Omega_1 \times \Omega_2$ and $u_i \in \Omega_i = [0, 8]$, $i = 1, 2$. An additional constraint is considered: $u_1 + u_2 \leq 8$. Now the action space is modified to be

$$\bar{\Omega} = \{ u \in \Omega \mid u_1 + u_2 - 8 \leq 0 \}.$$  

Fig. 2.2 shows that $\Omega$ is rectangular where the constraints have no coupling and $\bar{\Omega}$ is triangular where the constraints are coupled. For $\bar{\Omega}$, it is not possible to obtain separate action sets from which the players can take actions independently, i.e., $u_1 \in [0, 8 - u_2]$ and $u_2 \in [0, 8 - u_1]$. The action space $\bar{\Omega}$ is called *coupled* and this game is called a two-player Nash game with coupled constraints.
In this section we present results for Nash games with coupled constraints, i.e., coupled action sets. We call such a game as a coupled Nash game \(^2\).

In a coupled Nash game, the following coupled inequality constraints are considered:

\[
g_r(u) \leq 0, \quad r = 1, \ldots, R
\]

or \(g(u) \leq 0\) in a vector form with

\[
g(u) = \begin{bmatrix} g_1(u) \\ \vdots \\ g_R(u) \end{bmatrix}
\]

The associated coupled constraint set with each \(g_r(u)\) is denoted by

\[
\bar{\Omega}_r = \{u \in \Omega \mid g_r(u) \leq 0\}
\]

where \(\Omega\) is defined (2.17), i.e., \(\Omega = \Omega_1 \times \cdots \times \Omega_m\) with \(\Omega_i = [m_i, M_i]\).

Then the action space \(\bar{\Omega} \subset \mathbb{R}^m\) is coupled, defined as

\[
\bar{\Omega} = \bigcap_{r=1}^{R} \bar{\Omega}_r = \{u \in \Omega \mid g(u) \leq 0\}
\]

For every given \(u_{-i} \in \Omega_{-i}\), a projection action set is also defined for each \(i \in \mathcal{M}\),

\[
\hat{\Omega}_i(u_{-i}) = \{\xi \in \Omega_i \mid g(u_{-i}, \xi) \leq 0\}.
\]

The projection action set \(\hat{\Omega}_i(u_{-i})\) is the feasible action set under the given \(u_{-i}\) for each \(i \in \mathcal{M}\).

An individual cost function \(J_i : \bar{\Omega} \rightarrow \mathbb{R}\) for every \(i \in \mathcal{M}\) is defined satisfying Assumption 2.1. Thus the coupled Nash game is denoted by \(GAME(\mathcal{M}, \bar{\Omega}, J_i)\). A vector \(u = (u_{-i}, u_i)\) is called feasible if \(u \in \bar{\Omega}\). The concept of an NE solution is defined as follows.

\(^2\)Coupled Nash games are also called generalized Nash games (games with non-disjoint strategy sets), [30], games with coupled constraints, [69], social equilibria games or pseudo-Nash equilibria games [9,21].
Definition 2.4 (NE for coupled Nash games) A vector $u^* \in \Omega$ is called an NE solution of $\text{GAME}(\mathcal{M}, \hat{\Omega}_i, J_i)$ if for all $i \in \mathcal{M}$ and for every given $u^*_{-i}$,

$$J_i(u^*_{-i}, u^*_i) \leq J_i(u^*_{-i}, u_i), \ \forall \ u_i \in \hat{\Omega}_i(u^*_{-i}), \quad (2.34)$$

with $\hat{\Omega}_i(u^*_{-i})$ defined in (2.33).

We take a look at Example 2.3. Note that $(u^*_1, u^*_2) = \left(\frac{9}{14}, \frac{4}{7}\right) \in \Omega$ (see Fig. 2.3), where $(u^*_1, u^*_2)$ is the unique NE solution obtained in Example 2.1. It follows that $(u^*_1, u^*_2)$ is the NE solution of the associated Nash game with coupled constraints.

![Figure 2.3: Reaction curves in coupled Nash game: $r_1(u_2)$ and $r_2(u_1)$](image)

The following proposition (adapted from Theorem 4.4 in [12]) gives sufficient conditions for existence of an NE solution.

**Proposition 2.8 (Theorem 4.4, [12])** Let the action space $\Omega$ be a compact and convex subset of $\mathbb{R}^m$. Under Assumption 2.1, $\text{GAME}(\mathcal{M}, \hat{\Omega}_i, J_i)$ admits an NE solution.

Next we use the concept of NG cost functions presented in Section 2.2.1 to characterize the NE solution of $\text{GAME}(\mathcal{M}, \hat{\Omega}_i, J_i)$. 
The two-argument NG cost function $\tilde{J}(u; x)$ is defined as in (2.22),

$$\tilde{J}(u; x) := \sum_{i=1}^{m} J_i(u_{-i}, x_i).$$

Note that $\tilde{J}(u; x)$ is separable in its second argument $x$. Then with respect to $\tilde{J}(u; x)$, an NE solution $u^*$ of $GAME(M, \hat{\Omega}_i, J_i)$ satisfies

$$\tilde{J}(u^*; u^*) \leq \tilde{J}(u^*; x), \ \forall \ x \in \Omega, \ \text{with} \ g(u^*_i, x_i) \leq 0 \ \forall \ i \in M. \ \ (2.35)$$

This result can be obtained by using Definition 2.4, the definition of $\tilde{J}(u; x)$, (2.22) and the projection action set $\hat{\Omega}_i(u_{-i}), (2.33)$.

We also augment the coupled constraints $g(u)$ in (2.30) into an equivalent two-argument form, $\tilde{g}$,

$$\tilde{g}(u; x) = \sum_{i=1}^{m} g(u_{-i}, x_i), \ (2.36)$$

where $\tilde{g} = [\tilde{g}_1, \ldots, \tilde{g}_R]^T$ is with

$$\tilde{g}_r(u; x) = \sum_{i=1}^{m} g_r(u_{-i}, x_i), \ \forall \ r = 1, \ldots, R.$$  

If $x = u$, then it follows that

$$\tilde{g}(u; u) = \sum_{i=1}^{m} g(u) = m \times g(u). \ \ (2.37)$$

Note that $\tilde{g}(u; x)$ is also separable in its second argument $x$.

Intuitively, by introducing the NG cost function $\tilde{J}(u; x)$, the coupled Nash game is related to a constrained optimization problem that has a fixed-point solution. This methodology has been used for uncoupled Nash games in Section 2.2.1 where the associated optimization problem is unconstrained.

Particularly, the constrained optimization problem is a constrained minimization of $\tilde{J}(u; x)$, with respect to $x$, with constraints $\tilde{g}(u; x)$, (2.36). A solution $u^*$ of the constrained minimization of $\tilde{J}(u; x)$ satisfies

$$\tilde{J}(u^*; u^*) \leq \tilde{J}(u^*; x), \ \forall \ x \in \Omega, \ \text{with} \ \tilde{g}(u^*; x) \leq 0, \ \ (2.38)$$
with $\tilde{g}(u^*; u^*) \leq 0$. It has been proved by contradiction in [63] that $u^*$ is an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. We restate here for completeness.

Assume $u^*$ is not an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. It follows that for some $i \in \mathcal{M}$, there exists an $\tilde{x}_i \in \Omega_i$ with $g(u^*_{-i}, \tilde{x}_i) \leq 0$, such that

$$J_i(u^*_{-i}, \tilde{x}_i) < J_i(u^*_{-i}, u^*_i)$$

By adding a term $\sum_{j \in \mathcal{M}, j \neq i} J_j(u^*_{-j}, u^*_j)$ to both sides, the following inequality holds,

$$\tilde{J}(u^*; \tilde{x}) < \tilde{J}(u^*; u^*)$$

with $\tilde{x} := (u^*_{-i}, \tilde{x}_i) \in \Omega$. From $\tilde{g}(u^*; u^*) \leq 0$ and (2.37), it follows that $g(u^*_{-j}, u^*_j) \leq 0$, $\forall j \in \mathcal{M}$, such that

$$\tilde{g}(u^*; \tilde{x}) := \sum_{j \in \mathcal{M}, j \neq i} g(u^*_{-j}, u^*_j) + g(u^*_{-i}, \tilde{x}_i) \leq 0$$

It follows that there exists an $\tilde{x}$ such that

$$\tilde{J}(u^*; \tilde{x}) < \tilde{J}(u^*; u^*), \ \tilde{x} \in \Omega, \ \text{with} \ \tilde{g}(u^*; \tilde{x}) \leq 0,$$

which contradicts the hypothesis in (2.38). Thus $u^*$ is an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. We state this result as a proposition.

**Proposition 2.9** A solution $u^*$ of the constrained minimization of $\tilde{J}(u; x)$ satisfying (2.38) with $\tilde{g}(u^*; u^*) \leq 0$ is an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$.

Thus an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ can be found by solving the constrained minimization problem for $\tilde{J}(u; x)$ and searching for a fixed-point solution.

Note that for every given $u \in \Omega$, the constrained minimization problem is a standard constrained optimization problem introduced in Section 2.1, for which optimality conditions are given in Section 2.1 involving a set of Lagrange multipliers. The Lagrangian extension to the two-argument constrained optimization was proposed by Pavel in [63] and decomposition results were developed based on the Lagrangian extension.
Next we give a brief introduction of this methodology and restate some results for the purpose of further use in the following chapters. Readers are referred to [63] for detail proofs of the results.

A two-argument Lagrangian function $\tilde{L}$ is defined for $\tilde{J}$ and $\tilde{g}$, where

$$
\tilde{L}(u; x; \mu) = \tilde{J}(u; x) + \mu^T \tilde{g}(u; x),
$$

(2.39)

with $\mu = [\mu_1, \ldots, \mu_R]^T$ being the Lagrange multiplier vector.

The next result gives optimality conditions for an NE solution in terms of Lagrange multipliers, which is an extension to Propositions 2.3 and 2.4.

**Proposition 2.10** Consider $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. Let the action space $\hat{\Omega}$ be a compact and convex subset of $\mathbb{R}^m$ and Assumption 2.1 is satisfied.

(a) (Necessity): If $u$ is an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$, then there exists a unique vector $\mu^* \geq 0$ such that

$$
\nabla_x \tilde{L}(u; x; \mu^*) \bigg|_{x=u} = 0
$$

(2.40)

$$
\mu^*^T g(u) = 0
$$

(2.41)

where the notation “$|_{x=u}$” is defined in (2.26) and denotes finding a fixed-point solution.

(b) (Sufficiency): Let $u^*$ be a feasible vector together with a vector $\mu = [\mu_1, \ldots, \mu_R]^T$, such that $\mu \geq 0$ and $\mu^T g(u^*) = 0$. Assume that $u^*$ minimizes the Lagrangian function, $\tilde{L}$ (2.39), over $x \in \Omega$, as a fixed-point solution, i.e., $u = u^*$ satisfies

$$
u = \arg \min_{x \in \Omega} \tilde{L}(u; x; \mu),
$$

(2.42)

where minimization on RHS of (2.42) is done with respect to $x$ in $\tilde{L}$. Then $u^*$ is an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$.

Note that if $u$ is an inner NE solution, then $\mu^* = 0$ such that (2.40) is equivalent to (2.26), the necessary condition for an NE solution of uncoupled Nash games.
Remark 2.3 The Lagrangian optimality condition in Proposition 2.10 shows that $u^*$ is obtained by first minimizing the augmented Lagrangian function $\tilde{L}(u; x; \mu)$ with respect to the second argument $x$, which gives $x = \phi(u)$ for every given $u$. The next step involves finding a fixed-point solution $u^*$ of $\phi$ by setting $x = u$, i.e., solving $u = \phi(u)$. Note that $u^*$ thus obtained depends on $\mu$, $u^*(\mu)$. An optimal $\mu^*$ is achieved by solving

$$\mu^T g(u^*(\mu)) = 0 \text{ and } \mu^* \geq 0. \tag{2.43}$$

The obtained $u^*(\mu^*)$ is a solution of (2.38) and hence an NE solution of $GAME(M, \tilde{\Omega}, J_i)$. We call $(u^*(\mu^*), \mu^*)$ an NE solution-Lagrange multiplier pair.

Example 2.4 Consider the two-player Nash game presented in Example 2.3 with a coupled constraint. The corresponding NG cost function $\tilde{J}(u; x)$ is

$$\tilde{J}(u; x) = (2x_1^2 - 2x_1 - x_1u_2) + (x_2^2 - \frac{1}{2}x_2 - u_1x_2).$$

The augmented constraint is

$$\tilde{g}(u; x) = g(u_2, x_1) + g(u_1, x_2) = (x_1 + u_2 - 8) + (u_1 + x_2 - 8) \leq 0$$

Thus the Lagrangian function is obtained as

$$\tilde{L}(u; x; \mu) = \tilde{J}(u; x) + \mu \tilde{g}(u; x)$$

$$= (2x_1^2 - 2x_1 - x_1u_2) + (x_2^2 - \frac{1}{2}x_2 - u_1x_2) + \mu(x_1 + u_2 - 8 + u_1 + x_2 - 8)$$

To find an NE solution, $u$, and the corresponding Lagrange multiplier vector, $\mu$, we need solve one of the necessary conditions, (2.40) with $x_i = u_i, i = 1, 2$. It follows that

$$4u_1 = u_2 + 2 - \mu \text{ and } 2u_2 = u_1 + \frac{1}{2} - \mu$$

with $u$ and $\mu$ satisfying $\mu(u_1 + u_2 - 8) = 0$ and $\mu \geq 0$. Thus if $\mu = 0$, then $(u_1, u_2) = (\frac{9}{16}, \frac{4}{7})$, which is an NE solution. If $\mu \neq 0$, $(u_1, u_2) = (\frac{51}{16}, \frac{77}{16})$ with $\mu = -\frac{95}{16} < 0$ such that $(\frac{51}{16}, \frac{77}{16})$ is not an NE solution. \(\square\)
Next we present a decomposition result for the minimization of $\tilde{L}(u; x; \mu)$ (Theorem 3, [63]). The result shows that the minimum of $\tilde{L}(u; x; \mu)$ with respect to $x \in \Omega$ can be obtained by minimizing a set of one-argument Lagrangian functions.

We define a function called the dual cost function $D(\mu)$ as

$$D(\mu) := \tilde{L}(u^*; u^*; \mu),$$

where $u^*$ minimizes $\tilde{L}$ defined in (2.39) over $x \in \Omega$ as a fixed-point solution, i.e., $u = u^*$ satisfies

$$u = \arg\min_{x \in \Omega} \tilde{L}(u; x; \mu).$$

In a fixed-point notation, we can write $D(\mu)$ as

$$D(\mu) := \left[ \min_{x \in \Omega} \tilde{L}(u; x; \mu) \right] \left|_{\arg\min_{x \in \Omega} L = u} \right.,$$

where $\tilde{g}(u; \mu) \leq 0$. The decomposition result is restated as follows.

**Proposition 2.11 (Theorem 3, [63])** Consider GAME($\mathcal{M}, \hat{\Omega}, J_i$). Let the action space $\Omega$ be a compact and convex subset of $\mathbb{R}^m$ and Assumption 2.1 is satisfied. The associated dual cost function $D(\mu)$, (2.44), can be decomposed as

$$D(\mu) = \sum_{i=1}^{m} L_i(u^*_{-i}(\mu), u^*_i(\mu), \mu),$$

where

$$L_i(u_{-i}, x_i, \mu) = J_i(u_{-i}, x_i) + \mu^T \tilde{g}(u_{-i}, x_i)$$

and $u^*(\mu) = [u^*_i(\mu)] \in \Omega$ minimizes a set of $L_i$ defined in (2.46) over $x_i \in \Omega_i$ as a fixed-point solution, $\forall \ i \in \mathcal{M}$. In other words, $u_i = u^*_i(\mu)$ satisfies

$$u_i = \arg\min_{x_i \in \Omega_i} L_i(u_{-i}, x_i, \mu), \forall \ i \in \mathcal{M}.$$

Provided that the action space $\hat{\Omega}$ is compact and convex and Assumption 2.1 is satisfied, the NE solution-Lagrange multiplier pair $(u^*(\mu^*), \mu^*)$ can be obtained by solving a lower-level uncoupled Nash game with individual cost functions $L_i$, (2.46), so that $u^*(\mu)$ is obtained and a higher-level problem with respect to the Lagrange multiplier, i.e., solving (2.43).
2.3 Optical Network Setup

We consider an optical network that is defined by a set of optical fiber links \( \mathcal{L} = \{1, \ldots, L\} \) connecting optical switches. The optical switches allow channels in the network to be added, dropped or routed. They also provide the flexibility of channel power adjustments [2]. A link \( l \in \mathcal{L} \) is composed of \( N_l \) optical amplified spans. Each span includes an optical fiber followed by an optical amplifier. Fig. 2.4(a) depicts an optical network with a mesh topology.

![Diagram of mesh, link, and span](image)

Figure 2.4: A wavelength-routed optical network configuration: Mesh, link and span

A set of channels, \( \mathcal{M} = \{1, \ldots, m\} \), corresponding to a set of wavelengths, are transmitted across the network by intensity modulation and wavelength-multiplexing. We denote by \( \mathcal{M}_l \) the set of channels transmitted over link \( l \in \mathcal{L} \). Also, we denote by \( \mathcal{R}_i \), \( i \in \mathcal{M} \), the set of links from the associated Tx to the corresponding Rx that channel \( i \) uses in its optical route. We denote by \( u_i, n^0_i, p_i \) and \( n_i \) the channel signal power at Tx, the channel noise power at Tx, the channel signal power at Rx and the channel noise power at
Rx, respectively, for channel \( i \in \mathcal{M} \) (illustrated in Fig. 2.4(a)). We let \( u = [u_1, \ldots, u_m]^T \) denote the vector form. We also use \( u = (u_{-i}, u_i) \) with \( u_{-i} = [u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m]^T \). The end-to-end (Tx-to-Rx) optical signal-to-noise ratio (OSNR) for any channel \( i \in \mathcal{M} \), is defined as

\[
\text{OSNR}_i = \frac{p_i}{n_i}.
\]

(2.47)

The following provides the framework for modeling channel OSNR in optical networks.

An optical amplified span \( s \) on link \( l \) is composed of an optical fiber with loss coefficient, \( L_{l,s} \), which is wavelength independent, and an optical amplifier with gain \( G_{l,s,i} \). The optical amplifier introduces amplified spontaneous emission (ASE) noise, denoted by \( \text{ASE}_{l,s,i} \). Both the gain and ASE noise are wavelength dependent. Let \( p_{l,s,i} \) and \( n_{l,s,i} \) denote the channel signal and noise power, respectively, at the output of span \( s \) on link \( l \) (illustrated in Fig. 2.4(c)). When \( s = 0 \), we have

\[
\begin{align*}
 u_{l,i} & = p_{l,0,i}, \quad \text{(2.48a)} \\
 n_{l,i}^{in} & = n_{l,0,i}. \quad \text{(2.48b)}
\end{align*}
\]

The channel signal and noise power at the output of link \( l \), illustrated in Fig. 2.4(b), denoted by \( p_{l,i} \), \( n_{l,i}^{out} \), respectively, are taken to be the same as the channel signal and noise power at the output of span \( N_l \), i.e.,

\[
\begin{align*}
 p_{l,i} & = p_{l,N_l,i}, \quad \text{(2.49a)} \\
 n_{l,i}^{out} & = n_{l,N_l,i}. \quad \text{(2.49b)}
\end{align*}
\]

The channel signal and noise power at the input of link \( l \), denoted by \( u_{l,i} \) and \( n_{l,i}^{in} \), respectively, are identical to the signal and noise power at the output of the preceding link \( l' \) of link \( l \) on route \( \mathcal{R}_i \), i.e.,

\[
\begin{align*}
 u_{l,i} & = p_{l',i}, \quad \text{(2.50a)} \\
 n_{l,i}^{in} & = n_{l',i}^{out}. \quad \text{(2.50b)}
\end{align*}
\]
We define the span transmission at span \(s\) on link \(l\) for channel \(i\) as

\[
h_{l,s,i} = G_{l,s,i} L_{l,s}, \quad \forall \ s = 1, \ldots, N_l. \tag{2.51}
\]

**Assumption 2.2** ASE noise does not contribute to the amplifier gain saturation.

**Assumption 2.3** All spans on each link \(l\) have equal length and all optical amplifiers are operated in automatic power control (APC) mode with the same total power target \(P_0^l\) and have the same gain spectral shape [2, 46].

APC operation mode is a typical amplifier mode. In the APC operation mode, a constant total power is launched into each span of a link. Under Assumption 2.3, the following relation holds,

\[
\sum_{i \in \mathcal{M}_l} p_{l,s,i} = P_0^l, \quad \forall \ s = 1, \ldots, N_l, \quad \forall \ l \in \mathcal{L}. \tag{2.52}
\]

Specifically, at the output of link \(l\), (2.52) can be written as

\[
\sum_{i \in \mathcal{M}_l} p_{l,i} = P_0^l. \tag{2.53}
\]

The above condition is called the system constraint.

One benefit of keeping a constant total power, or a constant total launching power after each span, is the compensation of variations in fiber-span power loss across a link [27]. Moreover, if the total power target is selected to be below the threshold for nonlinear effects, an optimal gain distribution is achieved across the link after each span [46]. Under Assumption 2.3, by using a variable optical filter, the gain of an optical amplifier \(G_{l,s,i}\) can be decomposed as

\[
G_{l,s,i} = G_{l,i} f_{l,s}, \tag{2.54}
\]

where \(G_{l,i}\) is the corresponding gain value for channel \(i\) on the spectral shape of the optical amplifier on link \(l\) and \(f_{l,s}\) is the loss of the variable optical filter at span \(s\) on
link \( l \), adjusted to achieve the constant total power target \( P_l^0 \) \cite{27}. Then from (2.51) and (2.60) the span transmission at span \( s \) on link \( l \) for channel \( i \) can be written as

\[
h_{l,s,i} = G_{l,i} L_{l,s} f_{l,s}, \quad \forall \ s = 1, \ldots, N_l.
\]  

(2.55)

Based on (2.55), we define the **link transmission** on link \( l \) for channel \( i \) as

\[
T_{l,i} = \prod_{s=1}^{N_l} h_{l,s,i}, \quad \forall \ l = 1, \ldots, L.
\]  

(2.56)

The channel signal and noise power at the output of span \( s \) on link \( l \) are given as

\[
p_{l,s,i} = h_{l,s,i} p_{l,s-1,i}, \quad \text{(2.57a)}
\]

\[
n_{l,s,i} = h_{l,s,i} n_{l,s-1,i} + ASE_{l,s,i}. \quad \text{(2.57b)}
\]

Next we present channel OSNR model in optical networks developed based on (2.50), (2.53), (2.55) and (2.57).

### 2.3.1 OSNR Model

We first consider the OSNR model on a single point-to-point WDM link shown in Fig. 2.5. A set \( \mathcal{M} = \{1, \ldots, m\} \) of channels are transmitted over this link. The link consists of \( N \) cascaded optical amplified spans. We use the same notation as in the beginning of Section 2.3 but drop the link index.

![Figure 2.5: A point-to-point WDM fiber link](image)
Chapter 2. Background

Specifically, for span $s$, the fiber loss coefficient is denoted by $L_s$. We denote for each channel $i \in \mathcal{M}$, the gain of the optical amplifier after the $s^{th}$ span to be $G_{s,i}$ and the corresponding ASE noise to be $ASE_{s,i}$. We denote by $p_{s,i}$ and $n_{s,i}$ the optical signal and noise power of channel $i$ at the output of span $s$, respectively. Note that in a point-to-point WDM fiber link,

\[(s = 0) \quad u_i = p_{0,i}, \quad n_i^0 = n_{0,i}, \quad (2.58a)\]
\[(s = N) \quad p_i = p_{N,i}, \quad n_i = n_{N,i}. \quad (2.58b)\]

Under Assumption 2.3, we denote by $P^0$ the constant total power target, such that at the output of each $s^{th}$ span, (2.52) holds, i.e.,

\[
\sum_{j \in \mathcal{M}} p_{s,j} = P^0, \quad \forall s = 1, \ldots, N. \quad (2.59)
\]

Under Assumption 2.3, $G_{s,i}$ can be decomposed as

\[G_{s,i} = G_i f_s, \quad (2.60)\]

where $G_i$ is the corresponding gain value for channel $i$ on the spectral shape of the OA and $f_s$ is the loss of the variable optical filter. The span transmission on span $s$ and the link transmission for channel $i$ are given as

\[h_{s,i} = G_i L_s f_s, \quad (2.61)\]
\[T_i = \prod_{s=1}^{N} h_{s,i}, \quad \forall s = 1, \ldots, N. \quad (2.62)\]

The optical signal and noise power at the output of span $s$ are

\[p_{s,i} = h_{s,i} p_{s-1,i}, \quad (2.63a)\]
\[n_{s,i} = h_{s,i} n_{s-1,i} + ASE_{s,i}. \quad (2.63b)\]

The following result gives the OSNR model for a point-to-point WDM link developed based on (2.58), (2.59), (2.61), (2.62) and (2.63). An earlier similar simple model can be found in [27].
Lemma 2.1  Under Assumptions 2.2 and 2.3, in a point-to-point optical link, the optical signal and noise power at Rx are given as

\[ p_i = T_i u_i \quad \text{(2.64a)} \]
\[ n_i = T_i n_i^0 + T_i \sum_{s=1}^{N} \frac{ASE_{s,i}}{H_{s,i}} \quad \text{(2.64b)} \]

where \( T_i \) is defined in (2.61) and

\[ H_{s,i} := \prod_{q=1}^{s} h_{q,i} = \frac{P^0(G_i)^s}{\sum_{j \in M} (G_j)^s u_j} \quad \text{(2.65)} \]

with \( h_{s,i} \) defined in (2.62).

The channel OSNR at Rx is given as

\[ OSNR_i = \frac{u_i}{n_i^0 + \sum_{j \in M} \Gamma_{i,j} u_j} \quad \text{(2.66)} \]

where \( \Gamma = [\Gamma_{i,j}] \) is the link system matrix with

\[ \Gamma_{i,j} = \sum_{s=1}^{N} \frac{(G_j)^s ASE_{s,i}}{(G_i)^s P^0}. \quad \text{(2.67)} \]

Proof: The proof follows by developing propagation relations for both signal and noise.

Using (2.63) iteratively together with (2.58a) generates

\[ p_{s,i} = u_i \prod_{q=1}^{s} h_{q,i}, \quad \text{(2.68a)} \]
\[ n_{s,i} = n_i^0 \prod_{q=1}^{s} h_{q,i} + \prod_{q=1}^{s} h_{q,i} \sum_{r=1}^{s} ASE_{r,i} \prod_{v=1}^{r} h_{v,i}. \quad \text{(2.68b)} \]

Using (2.58b) together with (2.61), (2.64) follows. Next substituting both (2.68a) and (2.61) into (2.59) generates

\[ P^0 = \sum_{j \in M} u_j \prod_{q=1}^{s} h_{q,j} = \sum_{j \in M} u_j (G_j)^s \prod_{q=1}^{s} (L_q f_q), \]

from which we get

\[ \prod_{q=1}^{s} (L_q f_q) = \frac{P^0}{\sum_{j \in M} (G_j)^s u_j}. \]
Then together with (2.61) we get

\[
H_{s,i} = \prod_{q=1}^{s} h_{q,i} = \prod_{q=1}^{s} (G_i L_q f_q) = \frac{P^0(G_i)^s}{\sum_{j \in \mathcal{M}} (G_j)^s u_j}.
\] (2.69)

Then substituting both (2.69) and (2.64) into (2.47), the link OSNR becomes

\[
OSNR_i = \frac{p_i}{n_i} = \frac{u_i}{n_i^0 + \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{i,j} u_j} = \frac{u_i}{n_i^0 + \sum_{j \in \mathcal{M}} \left( \sum_{k=1}^{N} \frac{G_k}{G_i} \right)^s \frac{ASE_{s,i}}{P_0} u_j}.
\] (2.70)

Mathematically the OSNR model in (2.66) is similar to the wireless signal to interference ratio (SIR) model [5]. However, it has a richer system structure: \( \Gamma \) with \( \Gamma_{i,j} \) defined in (2.67) is a full matrix, with cross-coupling terms, non-zero diagonal elements and all elements dependent on specific network parameters (e.g., the gain and ASE noise of optical amplifiers). We can rewrite (2.66) as

\[
OSNR_i = \frac{u_i}{X_{-i} + \Gamma_{i,i} u_i},
\] (2.70)

where

\[
X_{-i} = n_i^0 + \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{i,j} u_j.
\] (2.71)

Hence OSNR is no longer a linear function of channel input power like SIR as in [5]. Furthermore, since optical amplifiers are often cascaded along the link, ASE noise accumulates over many amplifiers and degrades channel OSNR as the number of amplifiers increases, which is reflected by (2.66).

The network OSNR model can be derived in a similar way. The following result gives the OSNR model for optical networks, adapted from [62].

**Proposition 2.12 (Lemma 2, [62])** Under Assumptions 2.2 and 2.3, the network OSNR for channel \( i \) at Rx is given as

\[
OSNR_i = \frac{u_i}{n_i^0 + \sum_{j \in \mathcal{M}} \Gamma_{i,j} u_j},
\] (2.72)
with
\[ \Gamma_{i,j} = \sum_{l \in \mathcal{R}_i} \prod_{q \in \mathcal{R}_{i,l}} H_{q,N_q,j} \sum_{s=1}^{N_l} \left( \frac{G_{l,j}}{G_{l,i}} \right)_s \frac{ASE_{l,s,i}}{P_l^0}, \]
where \( \mathcal{R}_{l,i} \) is the set of links on \( \mathcal{R}_i \) before link \( l \) and \( H_{l,s,i} \) can be obtained recursively as
\[ H_{l,s,i} = \frac{P_l^0}{\sum_{j \in \mathcal{M}_l} \prod_{q \in \mathcal{R}_{j,l}} H_{q,N_q,j} \left( \frac{G_{l,j}}{G_{l,i}} \right)_s u_j}. \]

The network OSNR model presents the same form as the link OSNR model (2.1), but with a more complex system matrix \( \Gamma = [\Gamma_{i,j}] \), which we refer as network system matrix.

The OSNR model shows that as input power of one channel increases, thereby increasing its OSNR, the noise in the other channels increases, thus decreasing their OSNRs. Based on the OSNR model, we take channel input power as inputs to the optical network system and OSNR outputs and other network measurements as feedback signals to design power control algorithms such that channel OSNRs can be optimized.

### 2.3.2 Link Capacity Constraint

Recall that in Chapter 1 we have mentioned that in optical networks, nonlinear effects on optical fiber along links limit the total input power. After each span of optical fiber, the total power is influenced by the gain distribution of optical amplifiers [18]. Under Assumption 2.3, the same total power is launched into each span of a link. This leads to a uniform total power distribution along the link, which minimizes the effects of optical nonlinearity [46].

However, the optimal gain distribution of optical amplifiers is not applicable to limit nonlinear effects on the optical fiber at the beginning of each link, i.e., the fiber segment between Txs (or optical switches) and the first optical amplifier along the link. Since the effects of the nonlinearity can obviously be reduced by lowering the total launched power, the following condition is imposed as the link capacity constraint in optical networks:

\[ \sum_{i \in \mathcal{M}_l} u_{l,i} \leq P_l^0, \quad \forall \ l \in \mathcal{L} \]  

(2.73)
where \( u_{l,i} \) is the signal power of channel \( i \) at the input of link \( l \). When channel \( i \) is transmitted from its associated Tx directly to link \( l \), \( u_{l,i} = u_i \), where \( u_i \) is the signal power of channel \( i \) at Tx. Otherwise, due to the power propagation along links and among networks, \( u_{l,i} \) is a function of \( u_i \), or \( u \) where \( u \) is defined as the signal power vector of all channels at Tx. This is similar to capacity constraints in flow control [8]. However, capacity constraints in flow control are unchanged along links, while link power constraints in optical networks are propagated along links. The forgoing (2.73) can be rewritten as

\[
g_l(u) = \sum_{i \in \mathcal{M}_l} u_{l,i} - P^0_l \leq 0, \quad \forall l \in \mathcal{L}.
\] (2.74)

We are interested in the convexity of the constraint (2.74).

**Example 2.5** In a single point-to-point WDM fiber link (Fig. 2.5), channel \( i \) is transmitted from its associated Tx directly to the link. Then we can drop the link index and (2.74) is reduced to

\[
g(u) = \sum_{i \in \mathcal{M}} u_i - P^0 \leq 0.
\] (2.75)

The function \( g(u) \) in the single link case is always convex. \( \square \)

In optical networks, the coupled power constraints are propagated along links. The next example shows that convexity of (2.74) is no longer automatically ensured.

---

**Figure 2.6:** Three-link topology with three channels
Example 2.6 Consider the network shown in Fig. 2.6 with three links \((L = 3)\) and three channels \((m = 3)\) transmitted over links. Each link \(l\) has \(N_l\) spans and a total power target \(P^0_l\). We study the convexity of the constraint on the third link, \(g_3(u)\):

\[
g_3(u) = u_{3,1} + u_{3,2} - P^0_3.
\]

By using (2.50a), i.e., \(u_{l,i}\) is identical to \(p_{l',i}\) with \(l'\) being the preceding link of \(l\) on route \(R_i\), we have

\[
g_3(u) = p_{2,1} + p_{2,2} - P^0_3.
\]

Then by using the system constraint (2.53) on link 2, we get

\[
g_3(u) = p_{2,1} + p_{2,2} - P^0_3 = -p_{2,3} + (P^0_2 - P^0_3)
\]

By using (2.64a) in Lemma 2.1 and (2.50a), we have

\[
p_{2,3} = \frac{P^0_2 (G_{2,3})^{N_2} u_3}{(G_{1,1})^{N_1} u_1 + (G_{1,2})^{N_1} u_2},
\]

\[
u_{2,i} = \frac{P^0_1 (G_{1,i})^{N_i} u_i}{(G_{1,1})^{N_1} u_1 + (G_{1,2})^{N_1} u_2}, \quad i = 1, 2.
\]

It follows that \(g_3(u)\) is a complicated function of \(u\). In the Hessian matrix of \(g_3(u)\), the sign of \(\frac{\partial^2 g_3(u)}{\partial u_i^2}\) and \(\frac{\partial^2 g_3(u)}{\partial u_j^2}\) may be different. Thus the Hessian matrix \(\nabla^2 g_3(u)\) may not be positive semidefinite, which leads to the result that \(g_3(u)\) is no longer automatically convex.

\[\square\]

2.4 Summary

As a preface to the remaining chapters, this chapter provided a self-contained overview of mathematical optimization and Nash game theory. In addition some basic concepts in WDM optical networks, specifically network modeling, were introduced, forming a basis for the problems addressed in the next four chapters.
Chapter 3

Games in Point-to-Point Topologies

The OSNR optimization problem in optical networks belongs to a subclass of resource allocation in general communication networks [3,73]. In optical network systems, a signal over the same fiber link can be regarded as an interfering noise for others, which leads to OSNR degradation. A satisfactory OSNR at Rx for each channel may be achieved by regulating the input power per channel at Tx. As a first step towards solving the OSNR optimization problem in optical networks, in this chapter we study such a problem in a single point-to-point optical link. A Nash game played among channels is employed towards maximizing OSNR with the link capacity constraint. Sufficient conditions are derived for the existence and uniqueness of an NE solution. Two convergent iterative algorithms are developed towards finding the NE solution.

3.1 Introduction

Noncooperative game theory is a powerful tool to capture strategy interactions among self-interested players. The game-theoretic approach is employed to solve the OSNR optimization problem. A Nash game is played among channels. The interest in such an approach is motivated by the departure from the assumption of cooperation among channels. Particularly in large-scale optical networks, decisions are often made indepen-
dently by channels according to their own performance objectives. Moreover, a beneficial feature of a game formulation is that it leads itself to iterative distributed algorithms for the computation of an NE solution.

In optical networks cascaded amplified spans are present, as well as accumulation and self-generation of optical ASE noise. These specific realistic physical features have been considered in OSNR modeling. In order for game methods to be practical, they must incorporate realistic constraints of the underlying network systems. One such important constraint is the link capacity constraint. Such a coupled constraint arises because when channels share an optical fiber, the total power launched into the fiber needs to be restricted below the nonlinearity threshold \([2, 46]\). This constraint is satisfied at intermediary amplifier sites, which are operated in the APC mode, but not at Txs. Thus a Nash game with a coupled constraint is considered.

The remainder of this chapter is organized as follows. Section 3.2 presents the Nash game formulation for the OSNR optimization problem with the link capacity constraint. In Section 3.3, we prove the existence and uniqueness of an NE solution and characterize properties of the NE solution. Section 3.4 proposes two iterative distributed algorithms towards computing the NE solution. We provide simulation results in Section 3.5 and give a summary in Section 3.6.

### 3.2 Game Formulation

We study a point-to-point WDM fiber link shown in Fig. 2.5 in Chapter 2 in which reconfiguration is finished, i.e., channels will not be added or dropped while performing the optimization. A set \(\mathcal{M} = \{1, \ldots, m\}\) of channels are transmitted over the link. The link consists of \(N\) cascaded spans of optical fiber each followed by an OA. All OAs have the same gain spectral shape and the gain value for channel \(i\) is \(G_i\).

We denote \(u_i\) and \(n_i^0\) the signal power and noise power of channel \(i \in \mathcal{M}\) at Tx,
respectively. Similarly, we denote $p_i$ and $n_i$ the signal power and noise power of channel $i \in \mathcal{M}$ at Rx, respectively. Let $u = [u_1, \ldots, u_m]^T$ denote the vector form of the signal power at Tx. Equivalently, we write $u = (u_{-i}, u_i)$ with $u_{-i} = [u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m]^T$ in some context to represent the same vector $u$. The signal power at Tx is typically limited for every channel. That is, $u_i$ is in a bounded set $\Omega_i = [0, u_{\text{max}}]$ where $u_{\text{max}}$ is a positive constant. We use $\Omega$ to represent the Cartesian product of $\Omega_i$, $i \in \mathcal{M}$, $\Omega = \Omega_1 \times \cdots \times \Omega_m$. We also let $\Omega_{-i} = \Omega_1 \times \cdots \times \Omega_{i-1} \times \Omega_{i+1} \cdots \times \Omega_m$.

The OSNR of channel $i$ at Rx, $\text{OSNR}_i = \frac{p_i}{n_i}$, is given as in (2.66), (2.67), i.e.,

$$\text{OSNR}_i = \frac{u_i}{n_i^0 + \sum_{j \in \mathcal{M}} \Gamma_{i,j} u_j},$$

with

$$\Gamma_{i,j} = \sum_{s=1}^{N} \frac{(G_j)^{s} \text{ASE}_{s,i}}{(G_i)^{s} P^0},$$

where $\text{ASE}_{s,i}$ is the ASE noise power of channel $i$ at span $s$ and $P^0$ is the total power target of the link.

Equivalently, (3.1) can be rewritten as

$$\text{OSNR}_i = \frac{u_i}{X_{-i} + \Gamma_{i,i} u_i},$$

where

$$X_{-i} = n_i^0 + \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{i,j} u_j.$$

The OSNR model reflects that the signal power of one channel can be regarded as an interfering noise for others, which leads to OSNR degradation. Regulating the optical powers at Tx, i.e., allocating optical power as a resource among channels, aims to achieve a satisfactory OSNR for each channel at Rx.

Next we use the game-theoretic approach to solve an OSNR optimization problem from the user optimization point of view. The OSNR model helps develop the game formulation. Specifically we formulate the OSNR optimization problem as a Nash game, where channels are the players. The objective of each player is to maximize its utility.
related to individual channel OSNR. Each channel adjusts its power towards this goal in the presence of all other channels. The game settles at an equilibrium when no channel can improve its utility unilaterally.

Considering the link capacity constraint, the action space \( \bar{\Omega} \) is

\[
\bar{\Omega} = \left\{ u \in \Omega \mid \sum_{j \in M} u_j - P^0 \leq 0 \right\}, \tag{3.5}
\]

which is coupled in the sense that one player’s action affects the feasible action sets of the other players. The feasible action set for each channel \( i \) is the projection set

\[
\hat{\Omega}_i(u_{-i}) = \left\{ \xi \in \Omega_i \mid \sum_{j \in M, j \neq i} u_j + \xi - P^0 \leq 0 \right\}. \tag{3.6}
\]

An individual cost function \( J_i : \bar{\Omega} \to \mathbb{R} \) assigned to each channel \( i \in M \) is defined as the difference between a pricing function \( P_i : \bar{\Omega} \to \mathbb{R} \) and a utility function \( U_i : \bar{\Omega} \to \mathbb{R} \):

\[
J_i = P_i - U_i. \tag{3.7}
\]

Here the utility function \( U_i \) is chosen to be a logarithmic function of the associated channel’s OSNR (reasons are shown below):

\[
U_i = \beta_i \ln \left( 1 + \frac{a_i}{\text{OSNR}_i - \Gamma_{i,i}} \right), \quad \forall \ i \in M \tag{3.8}
\]

where \( \beta_i > 0 \) is a channel-defined parameter indicating the strength of the channel’s desire to maximize its OSNR and \( a_i > 0 \) is a scaling parameter for flexibility.

The utility function \( U_i \) defined in (3.8) is monotonically increasing in \( \text{OSNR}_i \). We illustrate the relationship in Fig. 3.1. Hence, maximizing the utility function \( U_i \) is related to maximizing \( \text{OSNR}_i \). From (3.3), the utility function can be equivalently written in terms of \( u_i \),

\[
U_i(u_{-i}, u_i) = \beta_i \ln \left( 1 + a_i \frac{u_i}{X_{-i}} \right), \quad \forall \ i \in M. \tag{3.9}
\]

It can be checked that the utility function \( U_i(u_{-i}, u_i) \) is twice continuously differentiable in its arguments, monotonically increasing and strictly concave in \( u_i \).
Some reasons for the choice of a logarithmic function (3.9) are as follows. A logarithmic function is analytically useful and is widely used as a utility function in flow control [37, 74] and power control [5] for general communication networks. In some cases, the logarithmic utility function is intimately associated with the concept of proportional fairness [37]. More specifically, $OSNR_i$ is strictly increasing with respect to $u_i$ and tends to $1/\Gamma_{ii}$ as $u_i$ tends to infinity. The OSNR model has a striking similarity with the wireless SIR model [5], but with a more general full system matrix $\Gamma$. Moreover, OSNR is no longer a linear function of its associated channel power. A direct logarithmic utility function of associated SIR as in the wireless case cannot be applied.

The pricing function consists of two terms: a linear pricing term and a regulation (penalty) term,

$$P_i(u_{-i}, u_i) = \alpha_i u_i + \frac{1}{P_0 - \sum_{j \in \mathcal{M}} u_j}, \quad \forall i \in \mathcal{M},$$

where $\alpha_i > 0$ is a pricing parameter determined by the system.

The linear pricing term is a linear function of the individual channel input signal power. It can be interpreted as the price a channel pays for using the system resources [72]. The linear pricing term reflects the fact that increasing one channel’s power degrades
the OSNR of all other channels. From the system performance point of view, the linear pricing term here is to limit the interferences of other channels caused by this channel and hence it improves the overall system performance [73].

The regulation term is constructed by considering the link capacity constraint. It penalizes any violation of the constraint in the following way: The regulation term tends to infinity when the total power approaches the total power target $P^0$, so the pricing function $P_i(u_{-i}, u_i)$ increases without bound. Hence the system resource is preserved by forcing all channels to decrease their input powers and indirectly satisfies the link capacity constraint.

Thus the $m$-player Nash game is defined in terms of the cost functions $J_i(u_{-i}, u_i)$, $i \in \mathcal{M}$, played within the action space $\hat{\Omega}$. This game is denoted by $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. Mathematically, in such a game, for each player $i$, given the actions of all other players, $u_{-i} \in \Omega_{-i}$, it independently solves a constrained optimization problem, i.e.,

$$\begin{align*}
\min & \quad J_i(u_{-i}, u_i) \\
\text{subject to} & \quad u_i \in \hat{\Omega}_i.
\end{align*}$$

### 3.3 NE Solution

A solution $u^*$ of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ is called an NE solution in the sense of Definition 2.4. If in addition, the solution is not on the boundary of the action space $\hat{\Omega}$, it is called an inner NE solution.

Notice that

$$\frac{1}{P^0 - \sum_{j \in \mathcal{M}} u_j} \to \infty \text{ as } u_i \to P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j.$$ 

Thus the points on the hyperplane $\{u \in \mathbb{R}^m | \sum_{j \in \mathcal{M}} u_j = P^0\}$ are not NE solutions of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. In addition, since $u_i = 0$ means channel $i$ is inactive in the link, an NE solution $u^*$ with a zero component, say, $u^*_1 = 0$, of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ implies that
Chapter 3. Games in Point-to-Point Topologies

channel 1 does not have any effect on the game. So in this case, the game is equivalent to the one in which \((m - 1)\) channels play and the NE solution to the \((m - 1)\)-player Nash game does not have a zero component. In this sense, we assume that any NE solution \(u^*\) to the \(m\)-player OSNR Nash game does not have zero components. Thus an NE solution \(u^*\) to the \(m\)-player OSNR Nash game is always inner. Part of the results in this section appeared in [55, 59].

3.3.1 Existence and Uniqueness

The following result provides sufficient conditions for existence and uniqueness of an inner NE solution.

**Theorem 3.1** \(\text{GAME}(\mathcal{M}, \hat{\Omega}_i, J_i)\) admits a unique inner NE solution \(u^*\), if \(\forall \, i \in \mathcal{M},\)

\[
\begin{align*}
  a_i &> \sum_{j \neq i, j \in \mathcal{M}} \Gamma_{i,j}, \quad (3.11) \\
  \beta_i &< \frac{\beta_{\min}}{\sum_{j \neq i, j \in \mathcal{M}} \frac{\Gamma_{i,j}}{a_j}}, \quad (3.12) \\
  \alpha_i &> \alpha_{\max} \sqrt{\sum_{j \neq i, j \in \mathcal{M}} \frac{\beta_j \Gamma_{i,j}}{a_j}}, \quad (3.13)
\end{align*}
\]

where \(\beta_{\min} = \min_{j \in \mathcal{M}} \beta_j\) and \(\alpha_{\max} = \max_{j \in \mathcal{M}} \alpha_j\).

**Proof:** \((Existence)\) The action space \(\tilde{\Omega}\) is a compact and convex set with a non-empty interior. Each cost function \(J_i(u_{-i}, u_i)\) is continuous and bounded and the first and second partial derivatives of \(J_i(u_{-i}, u_i)\) with respect to \(u_i\) are well defined on \(\tilde{\Omega}\) except the hyperplane \(\left\{ u \in \mathbb{R}^m | \sum_{j \in \mathcal{M}} u_j = P^0 \right\}\), given as

\[
\begin{align*}
  \frac{\partial J_i(u)}{\partial u_i} &= \alpha_i + \frac{1}{(P^0 - \sum_{j \in \mathcal{M}} u_j)^2} - \frac{\beta_i a_i}{X_{-i} + a_i u_i}, \quad \forall \, i \in \mathcal{M} \quad (3.14) \\
  \frac{\partial^2 J_i(u)}{\partial u_i^2} &= \frac{2}{(P^0 - \sum_{j \in \mathcal{M}} u_j)^2} + \frac{\beta_i a_i^2}{(X_{-i} + a_i u_i)^2}, \quad \forall \, i \in \mathcal{M}. \quad (3.15)
\end{align*}
\]

It follows that \(\frac{\partial^2 J_i(u)}{\partial u_i^2}\) is positive and therefore \(J_i(u_{-i}, u_i)\) is strictly convex in \(u_i\). Since each point on the hyperplane \(\left\{ u \in \mathbb{R}^m | \sum_{j \in \mathcal{M}} u_j = P^0 \right\}\) is not an NE solution, by Propo-
sition 2.8, the game with the action space $\hat{\Omega}$ admits an NE solution, which is the inner NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$.

Since an NE solution is inner, it follows from Proposition 2.7 that (2.21), $\frac{\partial J_i(u)}{\partial u_i} = 0$, $\forall i \in \mathcal{M}$, holds. The vector form is $\nabla J(u) = 0$, where $\nabla J(u)$ is defined in (2.3).

(Uniqueness) We prove the uniqueness by contradiction. Suppose $u^0 = [u^0_1, \ldots, u^0_m]^T$ and $u^1 = [u^1_1, \ldots, u^1_m]^T$ are both inner NE solutions with $u^0, u^1 \in \Omega$, $u^0 \neq u^1$. Then

\begin{align*}
\nabla J(u^0) &= 0 \\
\nabla J(u^1) &= 0
\end{align*}

By using (3.7), it follows that $\forall i \in \mathcal{M},$

\begin{align*}
n_i^0 + \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{i,j} u_j^0 + a_i u_i^0 &= \frac{\beta_i a_i}{1 + \frac{\beta_i a_i}{(P^0 - \sum_{j \in \mathcal{M}} u_j^0)^2}} \\
n_i^1 + \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{i,j} u_j^1 + a_i u_i^1 &= \frac{\alpha_i}{1 + \frac{\alpha_i}{(P^0 - \sum_{j \in \mathcal{M}} u_j^1)^2}}
\end{align*}

(3.17)

Let

$$\Delta u = u^0 - u^1 = [\Delta u_1, \ldots, \Delta u_m]^T.$$  

Define $u^\theta$ as a convex combination of $u^0$ and $u^1$

$$u^\theta := \theta u^0 + (1 - \theta) u^1, \ 0 < \theta < 1.$$  

(3.19)

Then differentiating $\nabla J(u^\theta)$ with respect to $\theta$ yields

\begin{align*}
\frac{d\nabla J(u^\theta)}{d\theta} &\quad \nabla^2 J(u^\theta) \frac{d(u^\theta)}{d\theta} = \nabla^2 J(u^\theta) \Delta u,
\end{align*}

(3.20)

where the notation $\nabla^2 J$ is the Jacobian of $\nabla J$ with

\begin{align*}
\frac{\partial^2 J_i(u^\theta)}{\partial u_i^2} &= \frac{2}{(P^0 - \sum_{j \in \mathcal{M}} u_j^\theta)^3} + \frac{\beta_i a_i^2}{(X_{-i}^\theta + a_i u_i^\theta)^3} \\
\frac{\partial^2 J_i(u^\theta)}{\partial u_i \partial u_j} &= \frac{2}{(P^0 - \sum_{j \in \mathcal{M}} u_j^\theta)^3} + \frac{\beta_i a_i \Gamma_{i,j}}{(X_{-i}^\theta + a_i u_i^\theta)^3}
\end{align*}

(3.21) (3.22)

where $X_{-i}^\theta = n_i^0 + \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{i,j} u_j^\theta$. 
By using (3.16), integrating (3.20) over \( \theta \) yields

\[
\int_0^1 \frac{d\nabla J(u^\theta)}{d\theta} d\theta = \nabla J(u^1) - \nabla J(u^0) = 0,
\]

(3.23)
i.e.,

\[
\int_0^1 (\nabla^2 J(u^\theta) \triangle u) d\theta = \triangle u \int_0^1 \nabla^2 J(u^\theta) d\theta = 0.
\]

(3.24)
It follows that \( \int_0^1 \nabla^2 J(u^\theta) d\theta = 0 \) if \( \triangle u \neq 0 \). The following proves the uniqueness by showing that under conditions (3.11), (3.12) and (3.13), \( \int_0^1 \nabla^2 J(u^\theta) d\theta \neq 0 \), which means that the only possible case for (3.24) to hold is \( \triangle u = 0 \), i.e., \( u^0 = u^1 \).

Denote the left-hand side (LHS) of (3.17) as

\[
x_i + y_i := n_i^0 + \sum_{j \in M, j \neq i} \Gamma_{i,j} u_j^0 + a_i u_i^0
\]

(3.25)
\[
y_i := n_i^1 + \sum_{j \in M, j \neq i} \Gamma_{i,j} u_j^1 + a_i u_i^1
\]

(3.26)
Then using (3.18) yields

\[
x_i = \sum_{j \in M, j \neq i} \Gamma_{i,j} \triangle u_j + a_i \triangle u_i
\]

(3.27)
Therefore, integrating (3.21) and using (3.27), (3.26), we obtain

\[
\int_0^1 \frac{\partial^2 J_i(u^\theta)}{\partial u_i^2} d\theta = \int_0^1 \frac{2}{(P^0 - \sum_{j \in M} u_j^0 - \theta \sum_{j \in M} \triangle u_j)^3} d\theta + \int_0^1 \frac{\beta_i a_i^2}{(x_i \theta + y_i)^2} d\theta
\]

\[
= \frac{(P^0 - \sum_{j \in M} u_j^0) + (P^0 - \sum_{j \in M} u_j^1)}{(P^0 - \sum_{j \in M} u_j^0)^2(P^0 - \sum_{j \in M} u_j^1)^2} + \frac{\beta_i a_i^2}{y_i(x_i + y_i)}
\]

By using (3.17), we obtain

\[
\int_0^1 \frac{\partial^2 J_i(u^\theta)}{\partial u_i^2} d\theta = \frac{(P^0 - \sum_{j \in M} u_j^0) + (P^0 - \sum_{j \in M} u_j^1)}{(P^0 - \sum_{j \in M} u_j^0)^2(P^0 - \sum_{j \in M} u_j^1)^2}
\]

\[
+ \frac{1}{\beta_i} (\alpha_i + \frac{1}{(P^0 - \sum_{j \in M} u_j^0)^2}) (\alpha_i + \frac{1}{(P^0 - \sum_{j \in M} u_j^1)^2})
\]

\[
= W_1 + \frac{1}{\beta_i} W_2 + \frac{\alpha_i}{\beta_i} W_3 + \frac{\alpha_i^2}{\beta_i}
\]

(3.28)
where $W_1$, $W_2$ and $W_3$ are positive constants with

$$W_1 = \frac{(P^0 - \sum_{j \in M} u_0^j) + (P^0 - \sum_{j \in M} u_1^j)}{(P^0 - \sum_{j \in M} u_0^j)^2(P^0 - \sum_{j \in M} u_1^j)^2}$$

$$W_2 = \frac{1}{(P^0 - \sum_{j \in M} u_0^j)^2(P^0 - \sum_{j \in M} u_1^j)^2}$$

$$W_3 = \frac{(P^0 - \sum_{j \in M} u_0^j)^2 + (P^0 - \sum_{j \in M} u_1^j)^2}{(P^0 - \sum_{j \in M} u_0^j)^2(P^0 - \sum_{j \in M} u_1^j)^2}$$

Similarly, integrating (3.22) yields

$$\int_0^1 \frac{\partial^2 J_i(u^\theta)}{\partial u_i \partial u_j} d\theta = W_1 + \Gamma_{i,j} \frac{\alpha_i \Gamma_{i,j}}{\beta_j} W_2 + \frac{\alpha_i^2 \Gamma_{i,j}}{\beta_j} W_3 + \frac{\alpha_i}{\beta_j}$$

Therefore, $\int_0^1 \nabla^2 J(u^\theta) d\theta$ can be written as

$$\int_0^1 \nabla^2 J(u^\theta) d\theta = W_1 \cdot B + W_2 \cdot C + W_3 \cdot D + E$$

where $B = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$, $C = \begin{bmatrix} \frac{1}{\beta_i} & \cdots & \frac{\Gamma_{i,m}}{\beta_{i,m}} \\ \vdots & \ddots & \vdots \\ \frac{\Gamma_{m,i}}{\beta_{m,m}} & \cdots & \frac{1}{\beta_m} \end{bmatrix}$, $D = \text{diag}(\alpha)$, $E = \text{diag}(\alpha)$.

The matrix $B$ is positive semi-definite. If (3.11) holds, the matrices $C$, $D$ and $E$ are strictly diagonally dominant. If (3.12) holds, i.e.,

$$\frac{1}{\beta_i} > \sum_{j \in M, j \neq i} \frac{\Gamma_{j,i}}{\beta_j a_j}$$

then $C^T$ is strictly diagonally dominant. Thus from Lemma A.1, the matrix $C$ is positive definite in the sense that its symmetric part, $\frac{1}{2}(C + C^T)$, is positive definite. Moreover,

$$\sum_{j \in M, j \neq i} \frac{\beta_i \Gamma_{j,i}}{\beta_j a_j} < \frac{\beta_i \Gamma_{j,i}}{\beta_j a_j} < 1$$

By using (3.13) and the foregoing, it follows that

$$\alpha_{\max} \sum_{j \in M, j \neq i} \frac{\beta_i \Gamma_{j,i}}{\beta_j a_j} < \alpha_i$$
and
\[
\frac{\alpha_i}{\beta_i} > \sum_{j \in \mathcal{M}, j \neq i} \frac{\alpha_j \Gamma_{j,i}}{\beta_j a_j}
\]
such that \( D^T \) is strictly diagonally dominant. Following similar arguments, it can be shown that
\[
\frac{\alpha_i^2}{\beta_i} > \sum_{j \in \mathcal{M}, j \neq i} \frac{\alpha_j^2 \Gamma_{j,i}}{\beta_j a_j},
\]
so \( E^T \) is strictly diagonally dominant. Thus from Lemma A.1, the matrices \( D \) and \( E \) are all positive definite. Recall the matrix \( B \) is positive semi-definite, then it follows that \( \int_0^1 \nabla^2 J(u) d\theta \) is full rank and thus \( \int_0^1 \nabla^2 J(u) d\theta \neq 0 \). From (3.24) and the foregoing, we have \( \Delta u = 0 \), i.e., \( u^0 = u^1 \). Thus the NE solution is unique. 

### 3.3.2 Pricing Strategy

Based on the results in Theorem 3.1, we discuss the parameter selection strategies. The link sets fixed channel prices and each channel decides its willingness \( \beta_i \) to obtain a satisfactory OSNR. From the necessary condition, \( \frac{\partial J_i(u)}{\partial u_i} = 0 \), we have
\[
\beta_i(OSNR_i) = \frac{\alpha_i}{a_i} (X_{-i} + a_i u_i(OSNR_i)) + \frac{X_{-i} + a_i u_i(OSNR_i)}{a_i (P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j - u_i(OSNR_i))^2}, \tag{3.31}
\]
where
\[
u_i(OSNR_i) = \frac{X_{-i}}{1/OSNR_i - \Gamma_{i,i}}.
\]
For a given lower OSNR bound \( \tilde{\gamma}_i \), we can show that if \( \beta_i \) is adjusted to satisfy the lower bound
\[
\beta_i > \frac{\alpha_i}{a_i} \frac{1 + (a_i - \Gamma_{i,i}) \tilde{\gamma}_i}{1 - \Gamma_{i,i} \tilde{\gamma}_i} X_{-i} + \frac{(1 - \Gamma_{i,i} \tilde{\gamma}_i)(1 + (a_i - \Gamma_{i,i}) \tilde{\gamma}_i)}{a_i (P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j - (P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j \Gamma_{i,i} + X_{-i} \tilde{\gamma}_i)^2 X_{-i}), \tag{3.32}
\]
each channel has \( OSNR_i > \tilde{\gamma}_i \).
3.3.3 Properties of the NE Solution

Theorem 3.1 shows the existence of a unique inner NE solution under certain conditions, but without giving an explicit expression of this solution. Next we characterize some properties of the NE solution, which directly lead to the development of two iterative algorithms towards finding the NE solution numerically.

Consider the necessary conditions \( \frac{\partial J_i(u)}{\partial u_i} = 0 \), i.e.,

\[
\alpha_i + \frac{1}{(P^0 - \sum_{j \in \mathcal{M}} u_j)^2} = \frac{\beta_i a_i}{X_i - a_i u_i}, \quad \forall i \in \mathcal{M}.
\]

(3.33)

The unique inner NE solution of \( \text{GAME}(\mathcal{M}, \hat{O}_i, J_i) \) satisfies (3.33).

Alternatively, the necessary condition can be written in terms of reaction functions. Recall that the action space \( \hat{\Omega} \) is compact and convex and the cost function \( J_i(u_{-i}, u_i) \) is twice continuously differentiable over \( \hat{\Omega} \) in its arguments and strictly convex in \( u_i \). A reaction function \( r_i \) for channel \( i \) is defined as

\[
r_i(u_{-i}) := \arg\min_{u_i \in \hat{\Omega}_i(u_{-i})} J_i(u_{-i}, u_i), \quad \forall i \in \mathcal{M}.
\]

(3.34)

Then it follows that for every given \( u_{-i} \), \( r_i(u_{-i}) \) is single-valued. Furthermore, by Berge’s Maximum Theorem (Theorem A.3), \( r_i(u_{-i}) \) is continuous. Thus, there is a unique point, \( u_i = r_i(u_{-i}) \) for every \( i \in \mathcal{M} \). In vector form, we have

\[
u = r(u),
\]

(3.35)

where \( r := [r_1(u_{-1}), \ldots, r_m(u_{-m})]^T \).

As discussed in Chapter 2, an NE solution \( u^* \in \hat{\Omega} \) satisfies

\[
u_i^* = r_i(u_{-i}^*), \quad \forall i \in \mathcal{M},
\]

(3.36)

or \( u^* = r(u^*) \). The reaction function \( r_i \) is highly nonlinear. So the NE solution is analytically intractable.

The reaction function is in fact the implicit solution of (3.33). Notice that over \( \hat{\Omega} \), LHS of (3.33) is a monotonically increasing function with respect to \( u_i \), while RHS of
(3.33) is monotonically decreasing. Therefore, for every \( i \in \mathcal{M} \), there exists a unique intersection between LHS and RHS of (3.33), which is the NE solution \( u^* \). Although the NE solution is highly nonlinear, we can derive more characteristics of the NE solution by studying the reaction function.

**Lemma 3.1** The reaction function \( r(u) \) has the following properties:

- \( r(u) \geq 0 \); (**non-negativity**)

- If \( u > u' \), then \( r(u) < r(u') \). (**monotonicity**)

**Proof:** Non-negativity is obvious from the definition of \( r(u) \). Suppose \( u \) and \( u' \) are two distinct vectors and satisfy \( u > u' \), i.e., \( u_i > u'_i \). We denote \( \tilde{u} = r(u) \) and \( \tilde{u}' = r(u') \).

For channel \( i \), using the necessary condition, (3.33), we have

\[
\frac{\partial J_i}{\partial u_i}(u_i, u'_i) \bigg|_{u_i = \tilde{u}_i} = \frac{\partial J_i}{\partial u_i}(u'_i, u'_i) \bigg|_{u_i = \tilde{u}'_i} = 0,
\]

i.e.,

\[
\alpha_i + \left( \frac{1}{(P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j - \tilde{u}_i)^2} \right) = \frac{\beta_i}{\tilde{u}_i + (\frac{n_0^i}{a_i} + \sum_{j \in \mathcal{M}, j \neq i} \frac{\Gamma_{i,j}}{a_i} u_j)}, \tag{3.37}
\]

\[
\alpha_i + \left( \frac{1}{(P^0 - \sum_{j \in \mathcal{M}, j \neq i} u'_j - \tilde{u}'_i)^2} \right) = \frac{\beta_i}{\tilde{u}'_i + (\frac{n_0^i}{a_i} + \sum_{j \in \mathcal{M}, j \neq i} \frac{\Gamma_{i,j}}{a_i} u'_j)}. \tag{3.38}
\]

Note that

\[
0 < P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j < P^0 - \sum_{j \in \mathcal{M}, j \neq i} u'_j
\]

and

\[
- \left( \frac{n_0^i}{a_i} + \sum_{j \in \mathcal{M}, j \neq i} \frac{\Gamma_{i,j}}{a_i} u_j \right) < - \left( \frac{n_0^i}{a_i} + \sum_{j \in \mathcal{M}, j \neq i} \frac{\Gamma_{i,j}}{a_i} u'_j \right) < 0.
\]

We plot the marginal utility figures for (3.37) and (3.38), i.e., LHS and RHS of (3.37) and (3.38) and note that \( \tilde{u}_i \) and \( \tilde{u}'_i \) are at the intersection of LHS and RHS. Fig. 3.2 shows that \( \tilde{u}_i < \tilde{u}'_i \). Therefore, \( \tilde{u} < \tilde{u}' \), and hence \( r(u) < r(u') \) from the definition. \[\blacksquare\]

It follows from Lemma 3.1 that the reaction function \( r(u) \) is not **standard**\(^1\).

\(^1\)The definition of “standard” is from [87]: a function \( I(p) \) is standard if for all \( p \geq 0 \), the following properties are satisfied: \( I(p) > 0 \); if \( p \geq p' \), then \( I(p) \geq I(p') \); for all \( \alpha > 1 \), \( \alpha I(p) > I(\alpha p) \).
3.4 Algorithms and Convergence Analysis

In this section, we develop iterative algorithms towards finding the unique inner NE solution and study their convergence properties. Distributed algorithms are preferred with less coordination among channels. Two types of iterative algorithms are presented next: a parallel update algorithm (PUA) and a gradient algorithm (GA). PUA is developed using the reaction functions and GA is developed based on the gradient descent method. Part of the results appeared in [57].

3.4.1 Parallel Update Algorithm (PUA)

In the parallel update algorithm, each channel $i$ updates its input power $u_i$ based on its associated reaction function $r_i(u_{-i})$. Let

$$\{u(n)\} := \{u(0), u(1), \ldots, u(n), \ldots\}$$
denote the sequence of channel input power vectors, where \( u(0) \) is the initial channel input power vector. Thus at each iteration time \((n + 1)\), channel input power \( u_i(n + 1) \) is updated by

\[
(PUA) \quad u_i(n + 1) = r_i(u_{-i}(n)), \; \forall i \in \mathcal{M},
\]

or equivalently in a vector form,

\[
u(n + 1) = r(u(n)).
\]

If the reaction function \( r(u) \) were standard, PUA would be the standard power control algorithm [87] and the convergence results in the synchronous case could be directly applied to PUA.

Consider two subsequences of \( \{u(n)\} \), \( \{u(2n)\} \) and \( \{u(2n + 1)\} \), defined as

\[
\{u(2n)\} := \{u(0), u(2), \ldots, u(2n), \ldots\},
\]

\[
\{u(2n + 1)\} := \{u(1), u(3), \ldots, u(2n + 1), \ldots\}.
\]

**Lemma 3.2** For \( GAME(\mathcal{M}, \bar{\Omega}_i, J_i) \) with \( \mathcal{M} = \{1, 2\} \). Let \( u(0) = 0 \). The subsequence \( \{u(2n)\} \) is strictly monotonically increasing and the subsequence \( \{u(2n + 1)\} \) is strictly monotonically decreasing. Each of them converges to fixed points \( u^1 \) and \( u^2 \), respectively.

Fig. 3.3 illustrates the convergence of sequence \( \{u_i(n)\} \) for channel \( i \).

**Proof:** We prove by induction. It is obvious \( u(1) > 0 = u(0) \) and \( u(2) > 0 = u(0) \), and we assume \( u(0) < u(2) < \cdots < u(2n) \). Since \( u(2n + 1) = r(u(2n)) \), then from the monotonicity property of \( r(u) \) in Lemma 3.1 we have

\[
u(1) > u(3) > \cdots > u(2n - 1) > u(2n + 1).
\]

Also from (3.40), we have \( u(2(n + 1)) = r(u(2n + 1)) \) and \( u(2n) = r(u(2n - 1)) \). From monotonicity of \( r(u) \) and (3.41), we have \( u(2(n + 1)) = r(u(2n + 1)) > r(u(2n - 1)) = u(2n) \). Hence \( \{u(2n)\} \) is strictly monotonically increasing. Similarly it follows \( u(2n + 3) < u(2n + 1) \), therefore \( \{u(2n + 1)\} \) is strictly monotonically decreasing. Furthermore, the
action space $\Omega$ is a compact set. Since a non-decreasing (or non-increasing) sequence has a limit point in a compact set, we have $\lim_{n \to \infty} u(2n) = u^1$ and $\lim_{n \to \infty} u(2n + 1) = u^2$. ■

**Lemma 3.3** For $GAME(\mathcal{M}, \hat{\mathcal{Q}}, J_i)$ with $\mathcal{M} = \{1, 2\}$. Let $u(0) = 0$. The sequence $\{u(n)\}$ converges to a limit point $u^0$, i.e., $\lim_{n \to \infty} u(n) = u^0$.

**Proof:** From Lemma 3.2, it suffices to show that $u^1 = u^2$. We use discrete-time Lyapunov stability theory [76] to prove the result. Construct a new system with the state vector defined as

$$q(n) := u(2n + 1) - u(2n). \tag{3.42}$$

Based on Lemma 3.1 and the fact that $u(0) = 0$, we have $q(n) > 0$, or component-wise, $q_i(n) > 0$. By using (3.40), we make the following notation:

$$Q(u(2n)) := q(n) = r(u(2n)) - u(2n). \tag{3.43}$$

It follows from Lemma 3.1 that $Q(x)$ is strictly decreasing with respect to $x$. Therefore,

$$q(n + 1) = Q(u(2n + 2)) = Q \circ r \circ r(u(2n)).$$
where (3.40) was used twice. We define \( G := Q \circ r \circ r \), thus \( q(n+1) = G(u(2n)) \). From (3.43), we can write \( q(n+1) \) as
\[
q(n+1) = G \circ Q^{-1}(q(n)).
\] (3.44)

Since \( Q(x) \) is strictly decreasing, \( Q^{-1} \) exists. We define \( F := G \circ Q^{-1} \), therefore
\[
q(n+1) := F(q(n)).
\] (3.45)

When \( q = 0 \), \( u \) is a constant vector. Hence, \( F(0) = 0 \).

Therefore we obtain a nonlinear discrete-time state space system (3.45), with
\[
q(0) = u(1) - u(0) \text{ and } F(0) = 0.
\] (3.46)

Now we turn to construct a Lyapunov function \( V(q(n)) : \mathbb{R}^m \to \mathbb{R} \), for the new system (3.45)-(3.46). We select \( V(q(n)) = q(n)^T q(n) \), which is positive definite.

Notice that using (3.42) yields
\[
q(n+1) - q(n) = u(2n+3) - u(2n+2) - u(2n+1) + u(2n)
\]
From Lemma 3.2, it follows that \( u(2n+3) - u(2n+1) < 0 \) and \( u(2n) - u(2n+2) < 0 \).

Therefore, \( q(n+1) - q(n) < 0 \), or component-wise, \( q_i(n+1) < q_i(n) \). Thus,
\[
\Delta V(q(n)) = V(q(n+1)) - V(q(n))
= q(n+1)^T q(n+1) - q(n)^T q(n)
= \sum_{i=1}^{m} [q_i^2(n+1) - q_i^2(n)] < 0
\]

Since \( V(q) \) is a continuous positive definite function and \( \Delta V(q) \) is negative definite, by discrete-time Lyapunov stability theory (pp765, [29]), \( V(q) \) is a strong Lyapunov function for system (3.45)-(3.46), and the origin of (3.45) is asymptotically stable. Hence \( q(n) \) converges to zero. From (3.42), this implies two subsequences \( \{u(2n)\} \) and \( \{u(2n+1)\} \) converge to the same limit point, i.e., \( u^1 = u^2 \). Therefore, \( \{u(n)\} \) converges to the limit point \( u^0 = u^1 = u^2 \).

Based on Lemma 3.2 and Lemma 3.3, the following result holds.
**Theorem 3.2** For GAME$(\mathcal{M}, \hat{\Omega}_i, J_i)$ with $\mathcal{M} = \{1, 2\}$. Let $u(0) = 0$. PUA converges to the NE solution $u^*$, i.e.,
\[
\lim_{n \to \infty} u(n) = u^*.
\] (3.47)

**Proof:** We take the limits of both sides of (3.40),
\[
\lim_{n \to \infty} u(n+1) = \lim_{n \to \infty} r(u(n)).
\] (3.48)

From Lemma 3.3, LHS of (3.48) equals $u^0$. From Lemma 3.1, $r(u)$ is a monotonic function. Moreover it is continuous, thus
\[
\lim_{n \to \infty} r(u(n)) = r\left(\lim_{n \to \infty} u(n)\right) = r(u^0).
\] (3.49)

By using (3.49), we obtain $u^0 = r(u^0)$. As the NE solution $u^*$ is unique and satisfies $u^* = r(u^*)$, we have $u^0 = u^*$. Thus \{u(n)\} converges to the unique NE solution $u^*$. ■

Even though PUA uses the nonlinear reaction function (3.35) and solving (3.35) depends on the powers of the other channels, the algorithm can be implemented in a distributed way. This is shown in the following revised first-order necessary condition of (3.33):
\[
\alpha_i + \frac{1}{\left(\frac{1}{P^0 - u_i(n+1) - \sum_{j \neq i} u_j(n)}\right)^2} = \frac{\beta_i a_i}{\left(\frac{1}{OSNR_i(n)} - \Gamma_i \alpha_i + a_i\right) u_i(n+1)}, \quad \forall \; i \in \mathcal{M}.
\] (3.50)

Each $u_i(n+1) = r_i(u_{-i}(n))$ is obtained by solving (3.50). By using the feedback information, $OSNR_i(n)$ and $u_T(n) = \sum_{j \in \mathcal{M}} u_j(n)$, we have
\[
u_i(n+1) = r_i(u_{-i}(n)) = r_i(OSNR_i(n), u_T(n)).
\]

Note that only associated channel power $u_i$, $OSNR_i$ and total power are needed. Fig. 3.4 depicts the algorithm acting on the link OSNR model.

**Remark 3.1** PUA may not converge when more than two channels ($m > 2$) exist. The intuitive reason is that each channel updates its optical power only based on instant costs.
Chapter 3. Games in Point-to-Point Topologies

Figure 3.4: Distributed PUA

and parameters, ignoring future implications of its action. Therefore, at some iterations
the total optical power of all other channels will exceed the target power $P^0$ when $m > 2$. Thus the nonlinear reaction function is no longer monotonic. To overcome the power fluctuations, we can use a relaxed PUA, in which a relaxation parameter is used to determine the step size that each channel takes towards finding the NE solution at each iteration step. For example, (3.39) can be modified as

$$ u_i(n + 1) = (1 - \mu_i)u_i(n) + \mu_i r_i(OSNR_i(n), u_T(n)),$$

where the coefficient $0 < \mu_i < 1$. It is promising that the relaxed PUA converges if $\mu_i$ is restricted.

PUA is developed by using the reaction function $r(u)$. It is restricted to implement when $m > 2$. In the next section we develop a gradient algorithm based on the gradient descent method [15] and we prove its convergence. The gradient algorithm can be used even when $m > 2$. For convenience, particularly for the simplification of convergence proof, we develop the algorithm in the continuous-time domain.
3.4.2 Gradient Algorithm (GA)

We consider a model where each channel uses a gradient algorithm to update its power, given as

\[ \dot{u}_i(t) = \frac{du_i}{dt} = -\mu \left( \frac{\partial J_i(u_{-i}, u_i)}{\partial u_i} \right), \quad \forall \ i \in \mathcal{M}, \tag{3.51} \]

where \( t \) is the continuous time variable and the coefficient \( \mu > 0 \).

Using (3.7) and (3.14), we rewrite (3.51) as

\[ (GA) \quad \dot{u}_i(t) = -\mu \left( \alpha_i + \frac{1}{(P^0 - \sum_{j \in \mathcal{M}} u_j(t))^2} \sum_{k=1}^{m} \tilde{\Gamma}_{i,k} u_k(t) \right) \tag{3.52} \]

where

\[ \tilde{\Gamma}_{i,k} = \begin{cases} \Gamma_{i,k}, & k \neq i, \\ a_i, & k = i. \end{cases} \]

Recalling the definition of OSNR, we can rewrite (3.52) as

\[ \dot{u}_i(t) = -\mu \left( \alpha_i + \frac{1}{(P^0 - \sum_{j \in \mathcal{M}} u_j(t))^2} \left( \frac{1}{\text{OSNR}_i(t)} + a_i - \Gamma_{i,i} \right) u_i(t) \right). \tag{3.53} \]

Hence, at each iteration, individual channels need only the sum of power of all channels and local measurements, namely, its own power and current OSNR level. Thus the iterative algorithm (3.53) is distributed.

Next we define a set \( \bar{\Omega}_\delta \), which is a subset of the action space \( \bar{\Omega} \). We slightly modify the bounded set \( \Omega_i = [0, u_{\max}] \) to be

\[ \tilde{\Omega}_i = [u_{\min}, u_{\max}], \quad \forall \ i \in \mathcal{M} \]

where \( 0 < u_{\min} < u_{\max} \). The Cartesian product of \( \tilde{\Omega}_i \) is given by \( \Omega_\delta \). The set \( \bar{\Omega}_\delta \) is defined as

\[ \bar{\Omega}_\delta = \left\{ u \in \Omega_\delta \mid \sum_{j \in \mathcal{M}} u_j - P^0 \leq 0 \right\}. \]

We set \( u_{\min} \) to be sufficiently small such that Theorem 3.1 holds, i.e., \( u^*_i > u_{\min} \) for all \( i \in \mathcal{M} \) and

\[ \frac{\partial J_i(u_{-i}, u_i)}{\partial u_i} \bigg|_{u_i = u_{\min}} < 0, \quad \forall \ u_{-i}, \ \forall \ i \in \mathcal{M}. \tag{3.54} \]
Notice that because of the capacity constraint, the following statement is always true:

$$\frac{\partial J_i(u_{-i}, u_i)}{\partial u_i} \bigg|_{u_i = P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j} > 0, \ \forall u_{-i}, \ \forall i \in \mathcal{M}. \quad (3.55)$$

Thus we claim that the set $\bar{\Omega}_\delta$ is invariant under the algorithm (3.52). That is, the trajectory remains inside $\bar{\Omega}_\delta$. It follows that the trajectory lies in $\bar{\Omega}$ if the initial state is in $\bar{\Omega}_\delta$. Moreover, the equilibrium of (3.52) in $\bar{\Omega}_\delta$ is the unique NE solution $u^*$ of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$.

The following result proves the convergence of algorithm (3.52).

**Theorem 3.3** For $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$, let the initial condition $u(0) \in \bar{\Omega}_\delta$. Then the update scheme (3.51) converges to the NE solution $u^*$ if

$$a_{\min} > \left(\frac{2P^0}{u_{\min}}\right)^2 \max_{i \in \mathcal{M}} \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{j,i}, \quad (3.56)$$

$$\beta_i > S \cdot \beta_{\max}, \quad (3.57)$$

where $a_{\min} = \min_{j \in \mathcal{M}} a_j$, $\beta_{\max} = \max_{j \in \mathcal{M}} \beta_j$ and

$$S = \left(\frac{2P^0}{u_{\min}}\right)^2 \frac{\max_{i \in \mathcal{M}} \sum_{j \in \mathcal{M}, j \neq i} \Gamma_{j,i}}{a_{\min}}. \quad (3.58)$$

**Proof:** Part I (construction of Lyapunov function): Let $\phi_i(u) := \dot{u}_i(t)$, where $\dot{u}_i(t)$ is defined in (3.51) or (3.52), and define a candidate Lyapunov function,

$$V(u) := \frac{1}{2} \sum_{i \in \mathcal{M}} \phi_i^2(u) \quad (3.59)$$

Note that $V(u)$ is restricted to the set, $\bar{\Omega}_\delta$. Because of the uniqueness of the NE solution $u^*$, $\phi_i(u) = 0, \forall i \in \mathcal{M}$, if and only if $u = u^*$. Therefore, $V(u)$ is strictly positive for all $u \neq u^*$.

From (3.52), taking the second derivative of $u_i$ with respect to time $t$, we obtain,

$$\ddot{u}_i(t) = -\mu \sum_j \left(\frac{2}{(P^0 - \sum_{j \in \mathcal{M}} u_j(t))^2} + \frac{\beta_i a_i \tilde{\Gamma}_{i,j}}{(\sum_{k \in \mathcal{M}} \tilde{\Gamma}_{i,k} u_k(t))^2}\right) \phi_j(u). \quad (3.60)$$
It is obvious that $\dot{\phi}_i(u) = \ddot{u}_i(t)$.

Taking the derivative of $V(u)$, (3.59), with respect to time $t$, we have

$$\dot{V}(u) = \sum_{i \in M} \phi_i(u) \cdot \dot{\phi}_i(u)$$

$$= -\mu \sum_{i \in M} \phi_i(u) \sum_{j \in M} \left( \frac{2}{(P^0 - \sum_{j \in M} u_j(t))^2} + \frac{\beta a_i \Gamma_{i,j}}{(\sum_{k \in M} \Gamma_{i,k} u_k(t))^2} \right) \phi_j(u),$$

where (3.60) is used. For simplicity, let

$$a(t) = \frac{2}{(P^0 - \sum_{j \in M} u_j(t))^2};$$

$$b_i(t) = \frac{\beta a_i}{(\sum_{k \in M} \Gamma_{i,k} u_k(t))^2};$$

$$\theta_{ij}(t) = a(t) + b_i(t) \Gamma_{i,j}.$$

Therefore, we rewrite (3.61) as

$$\dot{V}(u) = -\mu \sum_{i \in M} \sum_{j \in M} (a(t) + b_i(t) \Gamma_{i,j}) \phi_j(u) \phi_i(u)$$

$$= -\mu \sum_{i \in M} \phi_i(u) \sum_{j \in M} \theta_{ij}(t) \phi_j(u)$$

$$= -\mu \phi^T(u) \Theta(t) \phi(u),$$

where vector $\phi$ and matrix $\Theta(t)$ are defined as

$$\phi = [\phi_1, \phi_2, \ldots, \phi_m]^T$$

$$\Theta(t) = [\theta_{ij}(t)]_{m \times m},$$

i.e.,

$$\Theta(t) = \begin{bmatrix} a(t) & a(t) & \cdots & a(t) \\ a(t) & a(t) & \cdots & a(t) \\ \vdots & \vdots & \ddots & \vdots \\ a(t) & a(t) & \cdots & a(t) \end{bmatrix} + \begin{bmatrix} b_1(t) a_1 & b_1(t) \Gamma_{1,2} & \cdots & b_1(t) \Gamma_{1,m} \\ b_2(t) \Gamma_{2,1} & b_2(t) a_2 & \cdots & b_2(t) \Gamma_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_m(t) \Gamma_{m,1} & b_m(t) \Gamma_{m,2} & \cdots & b_m(t) a_m \end{bmatrix}. $$
It is obvious that if $\Theta(t)$ is uniformly positive definite, then $\dot{V}(u)$ in (3.65), is negative definite. Hence, the system is asymptotically stable by Lyapunov’s stability theorem [39]. This is what we will prove next.

**Part II (\(\Theta(t)\) uniformly positive definite):** We define $\Theta(t) := A(t) + B(t)$ with $A(t) = [a(t)]_{m \times m}$ and $B(t) = [b_i(t)\tilde{\Gamma}_{i,j}]_{m \times m}$, and note that the matrix $A(t)$ is uniformly positive semi-definite. Recall the sufficient conditions for the existence of a unique NE solution, (3.11), (3.12) and (3.13). By using (3.11), the matrix $B(t)$ is strictly diagonally dominant. It readily follows from Lemma A.1 that if $B(t)^T$ is strictly diagonally dominant, then the matrix $B(t)$ is uniformly positive definite, thus $\Theta(t)$ is uniformly positive definite.

**Part II.1 \((B(t)^T\) strictly diagonally dominant):** We show

$$b_i(t)a_i > \sum_{j \in M, j \neq i} b_j(t)\Gamma_{j,i} \tag{3.66}$$

Using the definition of $b_i(t)$ in (3.63), we rewrite (3.66) in the following way:

$$\frac{\beta_i a_i^2}{\bigg( \sum_{k \in M, k \neq i} \Gamma_{i,k} u_k + a_i u_i \bigg)^2} > \sum_{j \in M, j \neq i} \frac{\beta_j a_j \Gamma_{j,i}}{\bigg( \sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j \bigg)^2},$$

or,

$$\beta_i > \sum_{j \in M, j \neq i} \frac{\beta_j \Gamma_{j,i}}{a_j} \frac{\Gamma_{i,k} u_k + u_i}{\sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j} \bigg( \sum_{k \in M, k \neq i} \frac{\Gamma_{i,k} u_k + u_i}{\sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j} \bigg)^2 \tag{3.67}$$

Define

$$S_i := \sum_{j \in M, j \neq i} \frac{\Gamma_{j,i}}{a_j} \frac{\Gamma_{i,k} u_k + u_i}{\sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j} \bigg( \sum_{k \in M, k \neq i} \frac{\Gamma_{i,k} u_k + u_i}{\sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j} \bigg)^2 \tag{3.68}$$

and

$$\tilde{S}_i := \sum_{j \in M, j \neq i} \frac{\Gamma_{j,i}}{a_j} \frac{\Gamma_{i,k} u_k + u_i}{\sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j} \bigg( \sum_{k \in M, k \neq i} \frac{\Gamma_{i,k} u_k + u_i}{\sum_{k \in M, k \neq j} \Gamma_{j,k} u_k + a_j u_j} + 1 \bigg)^2 \tag{3.69}$$

So that a condition slightly stronger than (3.67) is $\beta_i > S_i \cdot \beta_{\max}$ and $S_i < 1$.

Hence, if $S_i < 1$ and $\beta_i > S_i \cdot \beta_{\max}$, then (3.67) or (3.66) holds, and $B(t)^T$ is strictly diagonally dominant.

**Part II.2 \((S_i < 1):** First we show that $S_i < 1$ under the sufficient condition, (3.56), in the statement.
Note that $u_i$ is constrained within $[u_{min}, P^0]$ (where $P^0$ is a weaker upper bound for $u_i$ compared to the link capacity constraint, $\sum_{j \in M} u_j \leq P^0$), and within the action space $\bar{\Omega}$ as well. Thus, we have
\[
\sum_{k \in M, k \neq i} \frac{\Gamma_{ik}}{a_k} u_k + u_i \leq \sum_{k \in M, k \neq i} \frac{\Gamma_{ik}}{a_k} + 1 \cdot \frac{P^0}{u_{min}},
\]
such that from (3.68), (3.69), it follows that
\[
S_i \leq \left( \frac{P^0}{u_{min}} \right)^2 \cdot \tilde{S}_i \tag{3.70}
\]
Recall (3.11), (3.12) and (3.13) again. Since $0 < \sum_{k \in M, k \neq i} \frac{\Gamma_{ik}}{a_i} < 1$, thus we have
\[
\frac{1}{2} < \sum_{k \in M, k \neq i} \frac{\Gamma_{ik}}{a_i} + 1 \leq 2.
\]
From (3.69) and the foregoing, it readily follows that
\[
\tilde{S}_i \leq 4 \cdot \sum_{j \in M, j \neq i} \frac{\Gamma_{j,i}}{a_j}, \tag{3.71}
\]
which used into the RHS of (3.70) yields
\[
S_i \leq \left( \frac{P^0}{u_{min}} \right)^2 \cdot 4 \cdot \sum_{j \in M, j \neq i} \frac{\Gamma_{j,i}}{a_j}, \quad \forall i \in M \tag{3.72}
\]
and moreover it implies that
\[
S_i \leq \left( \frac{P^0}{u_{min}} \right)^2 \cdot \frac{4}{\min_{j \in M, j \neq i} a_j} \sum_{j \in M, j \neq i} \Gamma_{j,i}, \quad \forall i \in M \tag{3.73}
\]
Now if (3.56) is true, i.e.,
\[
\min_j a_j > \left( \frac{2P^0}{u_{min}} \right)^2 \max_i \sum_{j \in M, j \neq i} \Gamma_{j,i},
\]
it follows that
\[
\min_{j \in M, j \neq i} a_j > \left( \frac{2P^0}{u_{min}} \right)^2 \sum_{j \in M, j \neq i} \Gamma_{j,i}, \tag{3.74}
\]
holds, or equivalently,
\[
\frac{1}{\min_{j \in M, j \neq i} a_j} \sum_{j \in M, j \neq i} \Gamma_{j,i} < \left( \frac{u_{min}}{2P^0} \right)^2. \tag{3.75}
\]
Therefore, from (3.73) and (3.75), it follows that $S_i < 1$.

**Part II.3 ($\beta_i > S_i \cdot \beta_{max}$):** We now turn to clarify the condition, $\beta_i > S_i \cdot \beta_{max}$. We show that if (3.57) and (3.58) hold, then $\beta_i > S_i \cdot \beta_{max}$.

From (3.73), we get an upper-bound for $S_i$,

$$S_i \leq \left( \frac{2P_0}{u_{min}} \right)^2 \max_i \sum_{j \in \mathcal{M}, j \neq i} \frac{\Gamma_{j,i}}{\min_{j \in \mathcal{M}, j \neq i} a_j}, \forall i \in \mathcal{M} \quad (3.76)$$

Recalling $S$, defined in (3.57), i.e.,

$$S = \left( \frac{2P_0}{u_{min}} \right)^2 \max_{i \in \mathcal{M}} \frac{\sum_{j \in \mathcal{M}, j \neq i} \Gamma_{j,i}}{a_{min}},$$

we see that $S_i \leq S$, $\forall i \in \mathcal{M}$. Therefore, if (3.57) holds, i.e., $\beta_i > S \cdot \beta_{max}$, it follows that $\beta_i > S_i \cdot \beta_{max}$, $\forall i \in \mathcal{M}$ holds also.

Finally, from (Part II.2) and (Part II.3), $S_i < 1$ and $\beta_i > S_i \cdot \beta_{max}, \forall i \in \mathcal{M}$, when (3.56) and (3.57) hold. It follows that (3.66) holds. Therefore, $B(t)^T$ is strictly diagonally dominant and thus $B(t)$ is uniformly positive definite. Therefore $\Theta(t)$ is uniformly positive definite and $\dot{V}(u)$ in (3.65) is negative definite. It readily follows that $\phi_i(u(t)) = \dot{u}_i(t) \rightarrow 0$, $\forall i \in \mathcal{M}$. This in turn implies that $u_i(t)$ converges to the unique NE solution point $u^*$.

A gradient algorithm has been studied for stability in a congestion control game [3, 4] where a special structure matrix is automatically uniformly positive definite in the proof. Contrastingly in our case, the corresponding matrix $\Theta(t)$ is more general. System dependent conditions are required for $\Theta$ to be positive definite.

Compared the conditions (3.56) and (3.57) for the stability of GA with the sufficient conditions (3.11)-(3.13) for the existence of a unique NE solution, it readily follows that $\beta_i$ is not only upper-bounded by (3.12), but also lower-bounded by (3.57).
3.5 Simulations

We simulate the iterative algorithms presented in the previous section in MATLAB. We investigate multiple channels on a single optical link shown in Fig. 2.5. The optical link has a total power capacity constraint $P^0 = 2.5$ mW (3.98 dBm). The associated system matrix $\Gamma$ for $m = 3$ is obtained as

$$\Gamma = \begin{bmatrix}
0.7463 & 0.7378 & 0.7293 \\
0.7451 & 0.7365 & 0.7281 \\
0.7438 & 0.7353 & 0.7269
\end{bmatrix} \times 10^{-4}.$$

For both iterative algorithms, the channel parameters in cost functions are selected satisfying both the sufficient conditions for existence of a unique NE solution, (3.11)-(3.13), and the conditions for stability of the update scheme, (3.56)-(3.57). Simulations are repeated with the following selected parameters: $\alpha = [0.001, 0.001, 0.001]$, $\beta = [1, 3, 2]$ and $a = [1, 1, 1]$.

Two distributed algorithms are implemented in simulation. At each iteration time $(n+1)$, $u_i(n+1)$ is obtained via PUA, i.e., via solving (3.50),

$$\alpha_i + \frac{1}{\left( P^0 - \sum_{j \in M, j \neq i} u_j(n) - u_i(n+1) \right)^2} = \frac{\beta_i a_i}{\left( \frac{1}{\text{OSNR}_i(n)} - \Gamma_{i,i} + a_i \right) u_i(n+1)}.$$

A discretized version of GA (3.52) is used to compute $u_i(n+1)$:

$$u_i(n+1) = u_i(n) - \mu \left( \alpha_i + \frac{1}{\left( P^0 - \sum_{j \in M} u_j(n) \right)^2} - \frac{\beta_i a_i}{\left( \frac{1}{\text{OSNR}_i(n)} + a_i - \Gamma_{i,i} \right) u_i(n)} \right),$$

where $\mu = 0.01$. In addition, the initial channel power value in GA are set as $u(0) = [0.116 0.121 0.126]$ (mW).

The Nash game with two channels is investigated first. Two channels compete for the power resources and the game settles down at the NE solution via PUA and GA. The evolutions in iterative time of channel input power, total power and OSNR are shown in Fig. 3.5, Fig. 3.6 and Fig. 3.7. Numerical simulations show wide fluctuations in PUA. The
fluctuations are largely avoided when using GA. Moreover, during the iterative process, the total power constraint is violated when using PUA, but not when using GA.

Next we investigate the Nash game with three channels \((m = 3)\). The input power, total input power and OSNR vs iteration time are plotted in Fig. 3.8, Fig. 3.9 and Fig. 3.10 respectively. During the iteration, the total input power does not exceed the
Figure 3.7: PUA and GA: channel OSNR

Figure 3.8: GA: channel power

link capacity constraint, in accordance with the fact that the trajectory lies in $\bar{\Omega}$ when the initial state is in $\bar{\Omega}_d$. Moreover, the total input power reaches the equilibrium point more quickly than the input powers and is smaller than the total power capacity constraint, $P^0$. These facts partly reflect that the extra term gives only a small effect on the cost rather than distorting the cost, however, it prevents channels from making full use of
input power resource. Simulation results showed above prove the performance of the algorithms and validate the analytic results.
3.6 Summary

In this chapter, a framework has been presented for a Nash game formulated towards OSNR optimization with the link capacity constraint in point-to-point WDM fiber links. The link capacity constraint introduced in Section 2.3.2 has been considered and imposed indirectly by adding a penalty term to each channel cost function. Conditions have been obtained for the existence of a unique inner NE solution to this Nash game. The NE solution is analytically intractable and developing iterative algorithms is not immediate. We have studied the properties of the NE solution and presented two distributed algorithms: PUA and GA, towards finding the unique inner NE solution. The algorithms use only local measurements and the current load of the link. The algorithms can be implemented via adjusting input powers, such that each channel maintains its individual OSNR optimally. Their convergence properties have been studied both theoretically and numerically.

It is worth mentioning that there exists an alternate approach to solve the constrained OSNR optimization problem in single point-to-point WDM fiber links. In [63] Lagrangian extension and decomposition results were developed and applied to convert a coupled Nash game into a lower-level uncoupled Nash game and a higher-level problem for pricing. This approach led to a solution that made full use of the input power resource. The extension of both approaches to general network configurations is of great interest.
Chapter 4

Games in Multi-link Topologies

OSNR optimization problems with coupled constraints in single point-to-point WDM fiber links has been studied in Chapter 3. The purpose of this chapter is to study OSNR optimization problems in multi-link topologies \(^1\). In network configurations, coupled constraints are propagated along fiber links and constraint functions become complicated (as shown in Example 2.6 in Chapter 2) from the end-to-end point of view. The theoretical results derived in Chapter 3 cannot be extended immediately, because it is difficult to build an appropriate cost function with a complicated penalty pricing term which is associated with constraints. Moreover, the convexity may not be automatically ensured. Thus it is intractable to apply Lagrangian extension and decomposition results presented in Chapter 2 directly.

In this chapter, we first show that convexity is satisfied in multi-link topologies with single-sink structures. Thus we formulate a Nash game among channels and apply Lagrangian extension and decomposition results for the computation of an NE solution. It has been shown in [63] that this formulation leads to a solution that made full use of the input power resource. From this point of view, we use this game formulation instead of the one proposed in Chapter 3. The general multi-link structure can be partitioned into

\(^1\)The multi-link topology studied in this thesis is a representative for selected paths extracted from a mesh configuration in which no closed-loops being formed by channel optical paths.
stages and each stage is with a single-sink structure. We formulate a partitioned Nash game for general multi-link topologies, composed of ladder-nested stage Nash games. Stage Nash games are played sequentially and solutions are interconnected. Based on Lagrangian extension and decomposition results, we develop a hierarchical iterative algorithm towards computing a solution of the partitioned Nash game.

4.1 Introduction

Work on games with coupled constraints has been going on for the past 40 years. From a computation point of view, the study of Nash equilibria of coupled Nash games presents severe analytical difficulties [23]. Theoretical results in standard optimization theory has received interest in the study of games. The work in [23] and its related work in [49, 50] have shown that an NE solution can be calculated by solving a variational inequality and the results express conditions in terms of variational inequalities and Karush-Kuhn-Tucker (KKT) conditions for the pseudo-gradient. Recently a procedural method is proposed in [63] for computing an NE solution, based on Lagrangian extension to a game theoretical framework. The work in [63] has been reviewed in Chapter 2. It has concluded that for convex constraints, a hierarchical decomposition may be performed and an NE solution is obtained by solving a lower-level Nash game with no coupled constraints and a higher-level problem for Lagrangian prices. The decomposition leads to an iterative hierarchical algorithm towards computing the NE solution.

Games in optical networks are considered in the class of $m$-player games with coupled utilities and constraints. However, as shown in Example 2.6 in Chapter 2, unlike capacity constraints in flow control [13, 38], coupled constraints in optical networks are propagated along links and convexity may not be satisfied. The non-convexity introduces additional complexities for analysis. In this chapter, we propose a partition approach. More precisely, we partition the general multi-link structure into stages and each stage is with a
single-sink structure where channels are dropped off at the output of a same link. Convexity is ensured in single-sink topologies. Then we formulate a partitioned Nash game for general multi-link topologies composed of ladder-nested stage Nash games. Based on Lagrangian extension and decomposition results, we discuss iterative computation of equilibria based on a three-level hierarchical algorithm, for channels, prices and stages. Part of this work appeared in [56, 58].

The remainder of this chapter is organized as follows. Section 4.2 studies games in single-sink multi-link topologies. Section 4.3 studies games in general multi-link topologies by defining a partitioned game with stages. An iterative hierarchical algorithm is proposed. Simulation results are provided in Section 4.4, followed by a summary in Section 4.5.

### 4.2 Games in Single-sink Multi-link Topologies

In this section, we use game-theoretic approaches to study the OSNR optimization problem in single-sink multi-link topologies. In a single-sink multi-link topology with a set \( \mathcal{L} = \{1, \ldots, L\} \) of links and a set \( \mathcal{M} = \{1, \ldots, m\} \) of channels, channels are added at different links, but dropped off at the output of the same link (e.g. Fig. 4.1). The input power of each channel \( i \in \mathcal{M} \) at Tx is bounded, i.e., \( u_i \in \Omega_i = [u_{min}, u_{max}] \) with \( u_{max} > u_{min} \geq 0 \). Let \( \Omega \) be the Cartesian product of \( \Omega_i \), \( i \in \mathcal{M} \) and \( \Omega_{-i} \) be the Cartesian product of \( \Omega_j \), \( j \neq i \), \( j \in \mathcal{M} \). We order the \( m \) channels in the following way: a set \( \mathcal{M}_1^1 = \{1, \ldots, m_1\} \) of \( m_1 \) channels are added into the network at the 1st link, a set \( \mathcal{M}_l^a = \left\{ \sum_{j=1}^{l-1} m_j + 1, \ldots, \sum_{j=1}^l m_j \right\} \) of \( m_l \) channels are added at link \( l \) and

\[
\sum_{l=1}^L m_l = m.
\]
For each channel $i \in \mathcal{M}$, $u_{i,j}$ and $p_{i,j}$ denote the signal power of channel $i$ at the input and output of link $l$, respectively. Each link $l$ has a total power target $P_0^l$.

### 4.2.1 Convexity Analysis

In this section we analyze the convexity of constraints in a class of multi-link topologies where only one sink exists.

**Example 4.1** Consider an example shown in Fig. 4.1, where three channels ($m = 3$) are transmitted over a single-sink 3-link network ($L = 3$).

![Figure 4.1: A single-sink multi-link topology](image)

Similar to Example 2.6 in Chapter 2, the link capacity constraints $g_i(u)$ can be achieved, given as

$$
g_1(u) = \sum_{i=1}^{2} u_{1,i} - P_1^0 = u_1 + u_2 - P_1^0 \leq 0$$

$$
g_2(u) = \sum_{i=1}^{4} u_{2,i} - P_2^0 = \sum_{i=1}^{2} p_{1,i} + u_3 + u_4 - P_2^0 = u_3 + u_4 - (P_2^0 - P_1^0) \leq 0$$

$$
g_3(u) = \sum_{i=1}^{6} u_{3,i} - P_3^0 = \sum_{i=1}^{4} p_{2,i} + u_5 + u_6 - P_3^0 = u_5 + u_6 - (P_3^0 - P_2^0) \leq 0,$$
where the system constraint (2.53) is used twice. Thus the coupled constraint set for link $l \in \mathcal{L}$, $\Omega_l = \{ u \in \Omega | g_l(u) \leq 0 \}$ is convex and $\Omega = \bigcap_{l \in \mathcal{L}} \Omega_l$ is also convex. $\square$

This result can be generalized for single-sink multi-link topologies with $m$ channels, presented in the following lemma.

**Lemma 4.1** Consider a single-sink multi-link topology with a set of links $\mathcal{L}$. For any link $l \in \mathcal{L}$, the coupled constraint set $\Omega_l$ is convex. The overall coupled action space $\bar{\Omega} = \bigcap_{l \in \mathcal{L}} \bar{\Omega}_l$ is a convex set.

**Proof:** We denote $\bar{m}_{l-1} = \sum_{j=1}^{l-1} m_j$ and $\bar{m}_l = \sum_{j=1}^l m_j$. Then

$$M^l = \{ \bar{m}_{l-1} + 1, \ldots, \bar{m}_{l-1} + m_l \}.$$ 

Thus we can rewrite the link capacity constraint (2.74) as

$$g_l(u) = \sum_{j=1}^{m_l} u_{m_{l-1} + j} - \Delta P_0^l \leq 0, \quad (4.1)$$

where $\Delta P_0^l$ is a modified power constraint related to the coupled propagated constraints. In fact $\Delta P_0^l$ is the remaining total power constraint after the propagated optical power of channels is considered. In Example 4.1, $\Delta P_0^2 = P_2^0 - P_1^0$. We can rewrite $g_l(u)$ as

$$g_l(u) = \sum_{j=m_{l-1}+1}^{\bar{m}_l} u_j - \Delta P_0^l \quad (4.2)$$

or

$$g_l(u) = e^T_{m_l} u - \Delta P_0^l \quad (4.3)$$

where

$$e^T_{m_l} = \left[ \begin{array}{c} 0^T_{\bar{m}_{l-1}} \ 1^T_{m_l} \ 0^T \end{array} \right] \quad (4.4)$$

is an $1 \times m$ vector with $1_{m_l}$, an $m_l \times 1$ all ones vector, and $0$, a vector of all 0 components with appropriate dimension. Thus $\bar{\Omega}_l$ is convex for all $l$, and so is $\bar{\Omega}$. $\square$
Remark 4.1 From (4.2), we can see that in the inequality constraint for link $l$, $g_l(u) \leq 0$, only channels in the set $M_l^a$ are coupled, i.e., channels that are added into the network at link $l$ are coupled. This is advantageous for a hierarchical decomposition, as we will show later.

4.2.2 Game Formulation

Based on the convexity result in Lemma 4.1, now we consider an $m$-player Nash game in a single-sink multi-link topology. The action space is

$$\bar{\Omega} = \bigcap_{l \in L} \{ u \in \Omega | g_l(u) \leq 0 \}.$$  

Let the action set of channel $i \in M$ be the projection of $\bar{\Omega}$ on channel $i$'s direction, namely,

$$\hat{\Omega}_i(u_{-i}) = \{ \xi \in \Omega_i | g_l(u_{-i}, \xi) \leq 0, \forall l \in L \}.$$

In the Nash game formulation, each channel $i \in M$ minimizes its individual cost function $J_i : \bar{\Omega} \rightarrow \mathbb{R}$,

$$J_i(u_{-i}, u_i) = \alpha_i u_i - \beta_i U_i(u_{-i}, u_i),$$  

(4.5)

with a utility $U_i$ indicating preference for better OSNR:

$$U_i(u_{-i}, u_i) = \ln \left( 1 + \frac{a_i}{\text{OSNR}_i - \Gamma_{i,i}} \right) = \ln \left( 1 + \frac{a_i u_i}{n_i^0 + \sum_{j \neq i} \Gamma_{i,j} u_j} \right),$$

where $\alpha_i, \beta_i > 0$ are weighting factors, $a_i > 0$ is a channel specific parameter and $\Gamma = [\Gamma_{i,j}]$ is the network system matrix being defined in (2.72). Unlike the channel cost functions in Chapter 3, here $J_i$ does not have an extra term to take the link capacity constraints into account.

We denote the game by $GAME(M, \hat{\Omega}_i, J_i)$, which is in the class of $m$-player Nash games with coupled utilities and coupled constraints. It follows from (4.5) that $J_i$ is continuously differentiable in its arguments and convex in $u_i$. Note that the overall coupled action space $\bar{\Omega}$ is compact. By Lemma 4.1, $\bar{\Omega}$ is convex as well. Then from
Proposition 2.8, $GAME(M, \hat{\Omega}_i, J_i)$ admits an NE solution. In the following, we use Lagrangian extension and decomposition results presented in Chapter 2 to compute an NE solution.

### 4.2.3 Decomposition

An NE solution of $GAME(M, \hat{\Omega}_i, J_i)$ can be computed by using the Lagrangian extension to the game-theoretic framework. We continue using the notations defined in Chapter 2 for the separable NG cost function $\tilde{J}(u; x)$, (2.22),

$$\tilde{J}(u; x) = \sum_{i=1}^{m} J_i(u_{-i}, x_i),$$

the separable augmented constraint $\tilde{g}(u; x)$, (2.36),

$$\tilde{g}(u; x) = \sum_{i=1}^{m} g(u_{-i}, x_i),$$

the augmented Lagrangian function $\tilde{L}(u; x; \mu)$, (2.39),

$$\tilde{L}(u; x; \mu) = \tilde{J}(u; x) + \mu^T \tilde{g}(u; x),$$

and the dual cost function $D(\mu)$, (2.44),

$$D(\mu) = \tilde{L}(u^*; u^*; \mu),$$

where $u^*$ is such that $u = u^*$ satisfies $u = \arg \min_{x \in \Omega} \tilde{L}(u; x; \mu)$. In a fixed-point notation,

$$D(\mu) = \left[ \min_{x \in \Omega} \tilde{L}(u; x; \mu) \right]_{\arg \min_{x \in \Omega} \tilde{L}=u},$$

where $\tilde{g}(u; u) \leq 0$.

Then $GAME(M, \hat{\Omega}_i, J_i)$ is related to a constrained minimization of $\tilde{J}$, (2.22), (2.36), with respect to the second argument $x$, that admits a fixed-point solution. Individual components of a solution $u^*$ to this constrained minimization constitute an NE solution to $GAME(M, \hat{\Omega}_i, J_i)$ in the sense of Definition 2.4. From Remark 2.3, we know that $u^*$ is
obtained by first minimizing $\tilde{L}(u; x; \mu)$ with respect to $x$. The next step involves finding a fixed-point solution which depends on $\mu$, $u^*(\mu)$. An NE solution-Lagrange multiplier pair $(u^*(\mu^*), \mu^*)$ is obtained if $\mu^T g(u^*(\mu^*)) = 0$ and $\mu^* \geq 0$.

The dual cost function $D(\mu)$ is related to the minimization of the associated Lagrangian function $\tilde{L}(u; x; \mu)$. For $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ with convex coupled constraints, Proposition 2.11 yields a decomposition into a lower-level Nash game with no coupled constraints and a higher-level problem for pricing. Particularly, Remark 4.1 states that only channels that are added into the network at link $l$, are coupled. This benefits the hierarchical decomposition for $D(\mu)$ by Corollary 2.11.

**Theorem 4.1** Consider $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ on a single-sink multi-link topology. A set, $\mathcal{M}_i^0$, of channels are added into network at link $l \in \mathcal{L}$, $\mathcal{L} = \{1, \ldots, L\}$ and $\mathcal{M} = \bigcup_{l \in \mathcal{L}} \mathcal{M}_i^0$. The dual cost function $D(\mu)$, (4.6), can be decomposed as

$$D(\mu) = \sum_{i=1}^{m} L_i(u_{-i}^*(\mu), u_i^*(\mu), \mu_{r(i)}) + \sum_{i=1}^{m} \sum_{l=1}^{L} \mu_l(e_{m_i, -i}^T u_{-i}^*(\mu) - \Delta P_0^l)$$

(4.7)

where $r(i)$ refers to link $r$ where channel $i$ is added, $e_{m_i, -i}$ is an $(m - 1)$ vector obtained by deleting the $i$th element of the $m$ vector $e_{m_i}$, and $u^*(\mu)$ is an NE solution to $GAME(\mathcal{M}, \Omega_i, L_i)$ where

$$\tilde{L}_i(u_{-i}, x_i, \mu_{r(i)}) = J_i(u_{-i}, x_i) + \mu_{r(i)} x_i.$$  

(4.8)

**Proof:** We order the $L$ links and $m$ channels in the same way as in the proof of Lemma 4.1. Then we rewrite the constraint function (4.3) in a two-argument form

$$g_i(u_{-i}, x_i) = e_{m_i, -i}^T u - \Delta P_0^l = \begin{cases} e_{m_i, -i}^T u_{-i} + x_i - \Delta P_0^l, & l = r(i) \\ e_{m_i, -i}^T u_{-i} - \Delta P_0^l, & l \neq r(i). \end{cases}$$

(4.9)

Since $J_i$ and $g_i$ are continuously differentiable and convex functions, by using Corollary 2.11, the dual cost function $D(\mu)$ can be decomposed as

$$D(\mu) = \sum_{i=1}^{m} L_i(u_{-i}^*(\mu), u_i^*(\mu), \mu)$$
where
\[ L_i(u_{-i}, x_i, \mu) = J_i(u_{-i}, x_i) + \mu^T g(u_{-i}, x_i). \]

Using (4.9) we see that \( \forall \ i \in \mathcal{M}, L_i \) is given as
\[
L_i = J_i(u_{-i}, x_i) + \mu r(i) x_i + \sum_{l \neq r(i)} \mu_l (e^T m_{l,-i} u_{-i} - \Delta P^0_l).
\] (4.10)

The following arguments are similar to those in Corollary 1 in [63]. We briefly outline here for completeness. We minimize RHS of (4.10) with respect to \( x_i \) and then solve for a fixed-point solution \( x_i = u_i \). Since only the first two terms in RHS of (4.10) depend on \( x_i \), isolating the terms without \( x_i \) yields
\[
D(\mu) = \sum_{i=1}^m \left[ \min_{x_i \in \Omega_i} \bar{L}_i(u_{-i}, x_i, \mu r(i)) \right]_{\arg \min_{x_i \in \Omega_i} L_i} + \sum_{i=1}^m \sum_{l=1}^L \mu_l (e^T m_{l,1} u_{-i} - \Delta P^0_l),
\] (4.11)
with \( \bar{L}_i \) as in (4.8). After minimization in each subproblem on RHS of (4.11), \( x^*_i \) is obtained as a function of \( u_{-i} \), namely,
\[ x^*_i = x_i(u_{-i}). \]

A fixed-point solution denoted by \( x_i = u_i \) can be obtained by setting \( x_i(u_{-i}) = u_i \) and solving simultaneously the \( m \) equations for a vector solution denoted by \( u^* \), or \( u^*(\mu) \) since it depends on \( \mu \). Substituting \( u^*(\mu) \) on RHS of (4.11) yields (4.7).

Theorem 4.1 leads to a hierarchical decomposition. An NE solution-Lagrange multiplier pair \((u^*(\mu^*), \mu^*)\) can be obtained by solving a lower-level uncoupled Nash game \( GAME(M, \Omega_i, \bar{L}_i) \) and a higher-level problem for coordination (i.e., setting price \( \mu \)).

The hierarchical game decomposition offers computational advantages. For a given price \( \mu \geq 0 \), the lower-level game with no coupled constraints admits an inner closed-form explicit NE solution. The following results give conditions for existence of the NE solution.
Proposition 4.1 For each given $\mu \geq 0$, GAME($\mathcal{M}, \Omega_i, \bar{L}_i$) admits an NE solution if $a_i$ in $\bar{L}_i$ (4.8) satisfies
\[ \sum_{j \neq i} \Gamma_{i,j} < a_i, \quad \forall \ i \in \mathcal{M}. \] (4.12)

The inner NE solution $u^*(\mu)$ is unique, given as
\[ u^*(\mu) = \tilde{\Gamma}^{-1} \cdot \tilde{b}(\mu), \]
where $\tilde{\Gamma} = [\tilde{\Gamma}_{i,j}]$ and $\tilde{b}(\mu) = [\tilde{b}_i(\mu)]$ are defined as
\[ \tilde{\Gamma}_{i,j} = \begin{cases} a_i & j = i, \\ \Gamma_{i,j} & j \neq i, \end{cases} \quad \tilde{b}_i(\mu) = \frac{a_i \beta_i}{\alpha_i + \mu r(i)} - n_i^0, \] (4.13)
with $\Gamma_{i,j}$ defined in (2.72).

Proof: The result can be proved by applying Theorem 2.6. We rewrite $\bar{L}_i$ (4.8) as
\[ \bar{L}_i(u_{-i}, u_i, \mu r(i)) = (\alpha_i + \mu r(i)) u_i - \beta_i \ln \left(1 + \frac{a_i u_i}{n_i^0 + \sum_{j \neq i} \Gamma_{i,j} u_j}\right), \quad \forall \ i \in \mathcal{M}. \] (4.14)
It can be seen that for any given $\mu \geq 0$, $\bar{L}_i$ is jointly continuous in all its arguments and
\[ \frac{\partial^2 \bar{L}_i}{\partial u_i^2} > 0, \quad \forall \ u_i \in \Omega_i. \]
It follows that $\bar{L}_i$ is strictly convex in $u_i$. Recall that for each $i \in \mathcal{M}$, $\Omega_i$ is a closed, bounded and convex subset of $\mathbb{R}$. Then by Theorem 2.6, GAME($\mathcal{M}, \Omega_i, \bar{L}_i$) admits an NE solution.

An inner NE solution can be found by solving the necessary conditions, $\frac{\partial \bar{L}_i}{\partial u_i} = 0$. From (4.14), we obtain
\[ a_i u_i^*(\mu) + \sum_{j \neq i} \Gamma_{i,j} u_j^*(\mu) = \frac{a_i \beta_i}{\alpha_i + \mu r(i)} - n_i^0, \quad \forall \ i \in \mathcal{M}. \]
Equivalently, in a matrix form, this is written as
\[ \tilde{\Gamma} \cdot u^*(\mu) = \tilde{b}(\mu), \] (4.15)
where matrix $\tilde{\Gamma}$ and vector $\tilde{b}(\mu)$ are defined in (4.13). Therefore a unique solution of (4.15) exists if the matrix $\tilde{\Gamma}$ is invertible. Notice that $\tilde{\Gamma}$ is a positive-entry matrix. If (4.12) holds, then $\tilde{\Gamma}$ is strictly diagonal dominant. From Gershgorin’s Theorem (Theorem A.1), it follows that $\tilde{\Gamma}$ is invertible and a unique solution of (4.15) exists,

$$u^*(\mu) = \tilde{\Gamma}^{-1} \cdot \tilde{b}(\mu),$$

which is an inner NE solution of $GAME(M, \Omega_i, \bar{L}_i)$.

\section*{4.3 Games in General Multi-link Topologies}

For general multi-link topologies, an NE solution of the OSNR Nash game formulated from the end-to-end point of view is intractable because of non-convexity of propagated constraints. The analysis results in Section 4.2 cannot be directly extended to general multi-link topologies.

Recall that the starting point for such an end-to-end game is that channel powers are independently adjustable only at Txs. In fact this physical restriction can be released. In optical networks, optical switches not only perform the function of selectively dropping some of wavelengths or adding selected wavelengths while letting others pass through, but also provide the flexibility of optical power adjustment for each channel \cite{2}.

In this section, we study the OSNR optimization problem in general multi-link topologies where channel powers are adjustable not only at Txs but also at the bifurcation nodes where channels are dropped. We partition the multi-link structure into stages depending on the nodes where channels are dropped (exit). Each stage has a single-sink structure and stages are ladder-nested \footnote{Nested structure is an information structure in which each player has an access to the information acquired by all his precedents, i.e., players situated closer to the beginning of the decision process. If the difference between the information available to a player and his closer precedent involves only his actions, then the structure is ladder-nested \cite{43}.}. The ladder-nested structure enables decomposition that in turn results in recursive procedure for its solution. Thus we turn to solve a set of convex
optimization problems which are related to the original single non-convex optimization problem.

In effect we define a partitioned Nash game where we exploit the single-sink structure of each stage and the ladder-nested form of the game. The game is played in the following way. Channels transmitted over the same stage compete for input power resource. Actions of competition are taken in a fixed order among stages and solutions are interconnected. In the following we will formulate a partitioned Nash game composed of ladder-nested stage Nash games. We will also discuss iterative computation of equilibria and develop a three-level hierarchical algorithm.

First, we make some new notations for each stage. A multi-link structure is partitioned into $K$ stages at the bifurcation nodes where channels are dropped. We let $\mathcal{K} = \{1, \ldots, K\}$ and for each stage $k \in \mathcal{K}$, we denote by $u_k = [u_{k,i}]$ and $p_k = [p_{k,i}]$ the input and output signal power vector, respectively. Similarly, we denote by $n_k^{in} = [n_{k,i}^{in}]$ and $n_k^{out} = [n_{k,i}^{out}]$ the input and output noise power vector, respectively. Sometimes we write $u_k$ as $u_k = (u_{k,-i}, u_{k,i})$ with the vector $u_{k,-i}$ obtained by deleting the $i^{th}$ element from $u_k$. Let $\gamma_k = [\gamma_{k,i}]$ denote the vector of channel adjustments at stage $k$ such that

$$u_{k,i} = \gamma_{k,i} p_{k',i},$$

where stage $k'$ is the stage precedent to stage $k$ for channel $i$. The adjustment parameter $\gamma_{k,i}$ is bounded within $[\gamma_{min}, \gamma_{max}]$. Each stage is composed of a set $\mathcal{L}_k = \{1, \ldots, L_k\} \subseteq \mathcal{L}$ of links. Each link $l \in \mathcal{L}_k$ has a constant total power target, $P_{0,k,l}$. A set $\mathcal{M}_k = \{1, \ldots, m_k\} \subseteq \mathcal{M}$ of channels are transmitted over stage $k$. Thus the OSNR for channel $i$, $i \in \mathcal{M}_k$, at the output of stage $k$, $k \in \mathcal{K}$, is defined as

$$\text{OSNR}_{k,i} = \frac{p_{k,i}}{n_{k,i}^{out}}.$$

We denote by $\mathcal{M}_{k,l}$ the set of channels transmitted over link $l$, $l \in \mathcal{L}_k$. Furthermore, we denote by $\mathcal{M}_{k,l}^a$ the set of channels added on stage $k$ at link $l$. Then the set of channels transmitted over the network $\mathcal{M} = \bigcup_{l \in \mathcal{L}_k} \mathcal{M}_{k,l}$. In the subsequent sections,
we sometimes use $r(i)$ instead of $l$, to refer to the associated link where channel $i$ is added. For channel $i$, we denote by $\hat{R}_i$ its optical route over stages, or collection of stages, from Tx to Rx, and $R_{k,i}$ its optical route on stage $k$. Specifically, we let $R_{k,l,i}$ be the optical route of channel $i$ before link $l$ on stage $k$, $l \in R_{k,i}$.

![Figure 4.2: Example: Three-link topology with three channels](image)

As an example, the multi-link topology in Fig. 4.2 can be partitioned into 2 stages ($K = 2$) at the bifurcation node OXC$_3$: the first stage is composed of link 1 and link 2; the second stage is composed of link 3. On stage 1, channels 1 and 3 are associated with links $r(1) = 1$, $r(3) = 2$, respectively. The defined optical routes of channel 1 are $\hat{R}_1 = \{1, 2\}$, $R_{1,1} = \{1, 2\}$, $R_{2,1} = \{3\}$ and $R_{1,2,1} = \{1\}$, respectively.

Based on partition, the OSNR of channel $i$ at the output of stage $k$, can be obtained from Proposition 2.12 as

$$OSNR_{k,i} = \frac{u_{k,i}}{n_{k,i}^{in} + \sum_{j \in M_k} \hat{\Gamma}_{k,i,j} u_{k,j}},$$

(4.16)

where $\hat{\Gamma}_k = [\hat{\Gamma}_{k,i,j}]$ is the stage system matrix, given as

$$\hat{\Gamma}_{k,i,j} = \sum_{l \in R_{k,i}} \prod_{q \in R_{k,l,i}} H_{q,N_q,j} \prod_{q \in R_{k,l,i}} H_{q,N_q,i} \Gamma_{l,i,j}$$

(4.17)
with
\[ \Gamma_{l,i,j} = \sum_{v=1}^{N_l} \left( \frac{G_{l,j}}{G_{l,i}} \right)^v \frac{ASE_{l,v,i}}{P_{k,l}^0}. \]
Based on (4.16), we define
\[ \frac{1}{\delta Q_{k,i}} := \frac{1}{OSNR_{k,i}} - \frac{1}{OSNR_{k',i}} = \sum_{j \in M_k} \tilde{\Gamma}_{k,j} \frac{u_{k,j}}{u_{k,i}} \]
where stage \( k' \) is the stage precedent to stage \( k \) for channel \( i \). It follows that \( \frac{1}{\delta Q_{k,i}} \) measures the OSNR degradation of channel \( i \) from stage \( k' \) to stage \( k \). Thus, instead of maximization of OSNR from Tx to Rx, we can consider minimization of OSNR degradation between stages.

### 4.3.1 Partitioned Game Formulation

Recall that by partitioning, each stage has a single-sink structure as in Section 4.2. At each stage \( k \), the set of channels is \( M_k = \{1, \ldots, m_k\} \). Let \( \tilde{\Omega}_k \) denote the action space on stage \( k \),
\[ \tilde{\Omega}_k = \bigcap_{l \in L_k} \left\{ u_k \in \prod_{i \in M_k} \Omega_i \mid g_{k,l}(u_k) \leq 0 \right\}, \]
and let the feasible action set of individual channel \( i \in M_k \) be the projection of \( \tilde{\Omega}_k \) on channel \( i \)'s direction, namely,
\[ \hat{\Omega}_{k,i}(u_{k,-i}) = \{ \xi \in \Omega_i \mid g_{k,l}(u_{k,-i}, \xi) \leq 0, \forall l \in L_k \}, \]
with
\[ g_{k,l}(u_k) = \sum_{i \in M_{k,l}} u_{k,i} - \Delta P^0_{k,l} = e^T_{m_{k,l}} \cdot u_k - \Delta P^0_{k,l} \leq 0, \forall l \in L_k, \]
where \( e^T_{m_{k,l}} \) is defined similarly as in (4.4) with all \( i \in M_k \) and \( \Delta P^0_{k,l} \) is a modified coupled power constraint on stage \( k \), defined as in Lemma 4.1. The associated vector form of (4.19) is denoted by
\[ g_k(u_k) = E_k^T u_k - \Delta P^0_k, \]
where $\Delta P_k^0$ is a vector with each element given as $\Delta P_{k,l}^0$ and $E_k$ is a matrix with each column vector $e_{k,l}$ given as $e_{k,l} = e_{m_{k,l}}$, (4.4). Since each stage has a single-sink structure, by Lemma 4.1, the linear coupled constraint function $g_{k,l}$ in (4.19) is convex and the overall coupled action space $\bar{\Omega}_k$ is also convex.

We formulate a partitioned Nash game, composed of lower-level ladder-nested stage Nash games. On stage $k$, a Nash game is played with each channel attempting to minimize its individual cost over $m_k$ channels. Games on stages are played in sequence and solutions are interconnected as we will show below.

We first specify channel individual cost function $J_{k,i}$ in the lower-level stage Nash game. Channel individual cost function $J_{k,i}$ has a similar form as (4.5), defined as a difference between a pricing function $\alpha_{k,i} u_{k,i}$ and a utility function. Differently the utility function here reflects the associated channel’s OSNR degradation. With the OSNR model (4.16) and new measure variable $\frac{1}{\delta Q_{k,i}}$ (4.18) for stage $k$, the new cost function $J_{k,i}$ is

$$J_{k,i} = \alpha_{k,i} u_{k,i} - \beta_{k,i} U_{k,i}, \forall i \in \mathcal{M}_k$$

(4.21)

with

$$U_{k,i} = \ln \left( 1 + \frac{a_{k,i}}{\delta Q_{k,i} - \hat{\Gamma}_{k,i}} \right),$$

(4.22)

where $\hat{\Gamma}_{k,i,j}$ is given as in (4.17). The utility function $U_{k,i}$ reflects the preference for lower OSNR degradation of channel $i$ on stage $k$. Substituting (4.18) into (4.21) yields

$$J_{k,i}(u_{k,-i}, u_{k,i}) = \alpha_{k,i} u_{k,i} - \beta_{k,i} \ln \left( 1 + a_{k,i} \frac{u_{k,i}}{\hat{X}_{k,-i}} \right),$$

(4.23)

where

$$\hat{X}_{k,-i} = \sum_{j \in \mathcal{M}_k, j \neq i} \hat{\Gamma}_{k,i,j} u_{k,j}.$$

(4.24)

It follows that $J_{k,i}(u_{k,-i}, u_{k,i})$ is continuously differentiable in its arguments and convex in $u_{k,i}$. We denote the lower-level stage Nash game by $GAME(\mathcal{M}_k, \hat{\Omega}_k, J_{k,i})$. Recall that the overall action space $\Omega_k$ is convex, then from Proposition 2.8, $GAME(\mathcal{M}_k, \hat{\Omega}_k, J_{k,i})$ admits an NE solution.
We exploit the partitioned Nash game composed of $K$ lower-level stage Nash games, given by $GAME(\mathcal{M}_k, \hat{\Omega}_{k,i}, J_{k,i})$. Each $GAME(\mathcal{M}_k, \hat{\Omega}_{k,i}, J_{k,i})$ is played by $m_k$ channels such that $\frac{1}{\delta q_{k,i}}$ is minimized for each channel $i$ on stage $k$. The lower-level stage Nash games are played sequentially (in a precedence order) with the interpretation that across all $K$ stages, $\sum_{k\in \hat{R}_i} J_{k,i}$ is related to the overall OSNR degradation for channel $i$. Solutions of all $GAME(\mathcal{M}_k, \hat{\Omega}_{k,i}, J_{k,i})$ are interconnected, explained as follows.

Let $u_k^* = [u_{k,i}^*]$ be an NE solution of $GAME(\mathcal{M}_k, \hat{\Omega}_{k,i}, J_{k,i})$. Recall that channel powers are adjustable at bifurcation nodes and $\gamma_{k,i}$ is the adjustable parameter for channel $i$ on stage $k$. Given the precedent actions of channel $i$, i.e., $u_{k',i}^*$ and $p_{k',i}^*$, the adjustment for channel $i$ on stage $k$ is obtained as

$$\gamma_{k,i}^* = \frac{u_{k,i}^*}{p_{k',i}^*}.$$

In a vector form, $\gamma_k^* = [\gamma_{k,i}^*]$, $i \in \mathcal{M}_k$, satisfies

$$u_k^* = \text{diag}(\gamma_k^*) \cdot p_{-k},$$

where $p_{-k}$ consists of the corresponding channel signal powers from the outputs of other stages for all $i \in \mathcal{M}_k$. Recall that each adjustment parameter is bounded within $[\gamma_{min}, \gamma_{max}]$. The partitioned Nash game admits a solution if each $\gamma_{k,i}^* \in [\gamma_{min}, \gamma_{max}]$.

Alternatively, $\gamma_k^*$ can be written as

$$\gamma_k^* = F_k(u_k^*, p_{-k}),$$

where $F_k$ is a one-to-one mapping.

It is worth mentioning that a similar partitioned game formulation was implemented for a single distributed optical link in [64]. The formulation for multi-link topologies is more complex. Firstly, at each stage, the coupled link capacity constraint has been considered. Secondly, the number of channels at different stages is variable while the number of channels at each $\gamma$-span ($\gamma$-stage) is identical in [64].
4.3.2 Iterative Hierarchical Algorithm

Since each stage has a single-sink structure, the decomposition results obtained in Section 4.2.3 for single-sink topologies can be directly extended to compute an NE solution of \( GAME(\mathcal{M}_k, \hat{\Omega}_{k,i}, J_{k,i}) \).

On each stage \( k \), \( GAME(\mathcal{M}_k, \hat{\Omega}_{k,i}, J_{k,i}) \) can be naturally decomposed into a lower-level Nash game with no coupled constraints and a higher-level stage problem for prices (Lagrange multipliers). By using Theorem 4.1, after decomposition, the lower-level Nash game is obtained with a modified cost function,

\[
L_{k,i}(u_{k,-i}, x_{k,i}, \mu_{k,r(i)}) = J_{k,i}(u_{k,-i}, x_{k,i}) + \mu_{k,r(i)}x_{k,i}, \quad i \in \mathcal{M}_k, \quad r(i) \in L_k,
\]  

where \( J_{k,i} \) is defined in (4.22) and an uncoupled action set \( \hat{\Omega}_i \). We denote this game by \( GAME(\mathcal{M}_k, \Omega_i, \bar{L}_{k,i}) \). We notice that \( \bar{L}_{k,i} \) has a similar form with \( \bar{L}_i \) (4.8). The following result characterizes the NE solution of \( GAME(\mathcal{M}_k, \Omega_i, \bar{L}_{k,i}) \) and can be easily derived from Proposition 4.1.

**Corollary 4.1** For each given \( \mu_k \geq 0 \), \( GAME(\mathcal{M}_k, \Omega_i, \bar{L}_{k,i}) \) admits an NE solution if \( a_{k,i} \) in \( \bar{L}_{k,i} \) (4.25) satisfies

\[
\sum_{j \in \mathcal{M}_k \setminus \{j \neq i\}} \hat{\Gamma}_{k,i,j} < a_{k,i}, \quad \forall \ i \in \mathcal{M}_k.
\]  

The inner NE solution \( u_k^*(\mu_k) \) is unique, given as

\[
u_k^*(\mu_k) = \tilde{\Gamma}_k^{-1} \cdot \text{diag}(a_k) \cdot \tilde{b}_k(\mu_k)
\]  

where \( \tilde{\Gamma}_k = [\tilde{\Gamma}_{k,i,j}] \) with

\[
\tilde{\Gamma}_{k,i,j} = \begin{cases} a_{k,i}, & j = i, \\ \hat{\Gamma}_{k,i,j}, & j \neq i, \end{cases}
\]

where \( \hat{\Gamma}_{k,i,j} \) is defined in (4.17) and \( \tilde{b}_k(\mu_k) = [\tilde{b}_{k,i}(\mu_k)] \) with

\[
\tilde{b}_{k,i}(\mu_k) = \frac{\beta_{k,i}}{\alpha_{k,i} + \mu_{k,r(i)}}.
\]
Remark 4.2 It can be seen that each $u_{k,i}^*(\mu_k)$ is continuously differentiable.

Next we propose an iterative hierarchical algorithm based on the NE solution (4.27) of the lower-level Nash game and coordination at the higher level.

For given $\mu_k \geq 0$, $u_{k}^*(\mu_k)$ in Corollary 4.1 is a solution to solve the Lagrangian optimality condition. Furthermore, $(u_{k}^*(\mu_k^*), \mu_k^*)$ is an NE solution-Lagrange multiplier pair if

$$\mu_k^T g_k(u_k^*(\mu_k^*)) = 0 \quad \text{and} \quad \mu_k^* \geq 0.$$ \hspace{1cm} (4.29)

By using (4.19), $\mu_k^* \geq 0$ can be obtained by solving

$$\mu_{k,l}^T \left( \sum_{j \in M_{k,l}} u_{k,j}^*(\mu_k^*) - \Delta P_{0,k,l} \right) = 0, \quad \forall \ l \in L_k.$$ \hspace{1cm} (4.30)

This shows that each stage acts like a coordinator setting the price at the optimal value $\mu_k^*$. Based on these, an iterative hierarchical adjustment algorithm is proposed for both stage pricing and channel power adjustment, which is recursive with respect to stages.

Channel Algorithm

On each stage $k$, based on price $\mu_k(t)$ from links, the following distributed iterative channel algorithm is used:

$$u_{k,i}(n+1) = \frac{\beta_{k,i}}{\alpha_{k,i} + \mu_{k,r(i)}(t)} - \left( \frac{1}{OSNR_{k,i}(n)} - \frac{1}{OSNR_{k',i}} - \hat{\Gamma}_{k,i,i} \right) \frac{u_{k,i}(n)}{a_{k,i}}$$ \hspace{1cm} (4.31)

where stage $k'$ is the precedent of stage $k$ and $OSNR_{k',i}$ is invariable during the channel iteration on stage $k$. Also note that the link price $\mu_{k,r(i)}(t)$ on stage $k$ is also constant during the channel iteration process. The algorithm (4.31) is distributed in the sense that each channel $i$ on stage $k$ updates its signal power $u_{k,i}$ based on the feedback information, i.e., its OSNR at the output of stage $k$, $OSNR_{k,i}$, and fixed parameters, $\mu_{k,r(i)}(t)$ and $OSNR_{k',i}$. 
Theorem 4.2  If for all \( i \in \mathcal{M}_k \), \( a_{k,i} \) in \( \bar{L}_{k,i} \) \( (4.25) \) are selected such that \( (4.26) \) is satisfied. Then for each given \( \mu_k \geq 0 \), channel algorithm \( (4.31) \) converges to the inner NE solution \( u_k^*(\mu_k) \).

Proof:  By Corollary 4.1, if on stage \( k \), \( a_{k,i} \) is selected satisfying \( (4.26) \), then for each given \( \mu_k \), an inner NE solution \( u_k^*(\mu_k) \), \( (4.27) \), exists and is unique in the sense of innerness. The solution for channel \( i \in \mathcal{M}_k \) is given as

\[
u_{k,i}^* = \frac{\beta_{k,i}}{\alpha_{k,i} + \mu_k r(i)} - \frac{1}{a_{k,i}} \sum_{j \in \mathcal{M}_k, j \neq i} \hat{\Gamma}_{k,i,j} u_{k,j}^*.
\]

(4.32)

The rest of the proof follows directly from the proof of Lemma 4 in [64]. We state here for completeness. Let

\[
ed_{k,i}(n) := u_{k,i}(n) - u_{k,i}^*(\mu_k).
\]

The corresponding vector form is \( \ned_k(n) = [\ldots, \ned_{k,i}(n), \ldots]^T \). We also define

\[
\|\ned_k(n)\|_\infty := \max_{i \in \mathcal{M}_k} |\ned_{k,i}(n)|.
\]

We can show that under condition \( (4.26) \), \( \|\ned_k(n+1)\|_\infty \leq C_0 \|\ned_k(n)\|_\infty \), where \( 0 \leq C_0 < 1 \) and \( \|\ned_k(n)\|_\infty \leq C_0^0 \|\ned_k(0)\|_\infty \), such that the sequence \( \{\ned_k(n)\} \) converges to 0. Therefore channel algorithm \( (4.31) \) converges to the inner NE solution \( u_k^*(\mu_k) \). \( \blacksquare \)

Link Algorithm

First we define a function based on \( (4.20) \),

\[
f_k(\mu_k) := -g_k(u_k^*(\mu_k)) = \Delta P_k^0 - E_k^T u_k^*(\mu_k).
\]

(4.33)

Recall \( (4.29) \) and \( (4.30) \). The optimal link prices \( \mu_k^* \) should satisfy

\[
f_k(\mu_k^*) \geq 0 \quad \text{and} \quad \mu_k^T f_k(\mu_k^*) = 0.
\]

(4.34)

We now select a constant \( \mu_{\max} > 0 \) and construct a set

\[
S_k := \{\mu_k \geq 0 | \mu_{k,l} \leq \mu_{\max}, \forall l \in \mathcal{L}_k\}.
\]
We come to find $\mu_k^* \in S_k$ such that (4.34) holds. This problem is a variational inequality problem introduced in Section A.7, denoted by $\text{VI}(S_k, f_k)$ in our setup.

Equivalently, the objective of $\text{VI}(S_k, f_k)$ is to determine a vector $\mu_k^* \in S_k$ such that

$$
(\mu_k - \mu_k^*)^T f_k(\mu_k^*) \geq 0, \forall \mu_k \in S_k.
$$

(4.35)

It follows from the definition of $S_k$ that $S_k$ is nonempty, compact and convex. The function $f_k$ is continuously differentiable on $S_k$. Thus $f_k$ is locally Lipschitz continuous on $S_k$ (Proposition A.1). Since $S_k$ is compact and convex, it follows immediately that $f_k$ is Lipschitz continuous on $S_k$ with Lipschitz constant $C_k$. That is

$$
\|f_k(\mu) - f_k(\nu)\| \leq C_k \|\mu - \nu\|, \forall \mu, \nu \in S_k.
$$

(4.36)

Moreover, $f_k$ is monotone on $S_k$ under certain conditions.

Lemma 4.2 If $a_{k,i}$ in $\bar{L}_{k,i}$ (4.25) satisfies

$$
a_{k,i} \geq 2 \max\{\sum_{j \neq i} \hat{\Gamma}_{k,i,j}, \sum_{j \neq i} \hat{\Gamma}_{k,j,i}\}, \forall i \in M_k,
$$

(4.37)

then the following condition holds,

$$
(\text{monotonicity}) \quad (\mu - \nu)^T (f_k(\mu) - f_k(\nu)) \geq 0, \forall \mu, \nu \in S_k.
$$

(4.38)

Remark 4.3 Obviously (4.37) is stronger than (4.26).

Proof: We start the proof with finding an explicit form of $f_k(\mu_k)$. From (4.27), the NE solution is

$$
u_k^*(\mu_k) = M_k^{-1}(a_k) \cdot \tilde{b}_k(\mu_k)
$$

with $M_k(a_k)$ defined as $M_k(a_k) = \text{diag}(\frac{1}{a_k}) \cdot \hat{\Gamma}_k$. It follows from the structure of $\tilde{\Gamma}_k$ that $M_k(a_k)$ is with diagonal elements being 1 and off-diagonal elements $\frac{\hat{\Gamma}_{k,i,j}}{a_{k,i}}$. We write $M_k(a_k)$ as $M_k(a_k) = I - A_k(a_k)$, where $A_k(a_k)$ is with diagonal elements being 0 and
off-diagonal elements \( \frac{\tilde{f}_{k,i}}{a_{k,i}} \). It follows directly from (4.37) that the induced norms of \( A_k(a_k) \) satisfy \( \|A_k(a_k)\|_1 \leq \frac{1}{2} \) and \( \|A_k(a_k)\|_\infty \leq \frac{1}{2} \). Since

\[
\|A_k(a_k)\|_2^2 \leq \|A_k(a_k)\|_1 \cdot \|A_k(a_k)\|_\infty,
\]

we also have \( \|A_k(a_k)\|_2 \leq \frac{1}{2} \). For simplicity we write \( \|A_k(a_k)\| \leq \frac{1}{2} \). We obtain \( \rho(A_k(a_k)) \leq \|A_k(a_k)\| < 1 \). It follows from Theorem A.2 that

\[
(I - A_k(a_k))^{-1} = \sum_{q=0}^{\infty} (A_k(a_k))^q
\]

exists. We denote \( F_k(a_k) = \sum_{q=1}^{\infty} (A_k(a_k))^q \). Together with \( \|A_k(a_k)\| \leq \frac{1}{2} \), we obtain for any induced norm of \( F_k(a_k) \),

\[
\|F_k(a_k)\| \leq \sum_{q=1}^{\infty} \|A_k(a_k)\|^q = \frac{\|A_k(a_k)\|}{1 - \|A_k(a_k)\|} \leq 1.
\]

Therefore, together with \( M_k^{-1}(a_k) = I + F_k(a_k) \), it follows from \( \|F_k(a_k)\|_1 \leq 1 \) and \( \|F_k(a_k)\|_\infty \leq 1 \) that both \( M_k^{-1}(a_k) \) and \( (M_k^{-1}(a_k))^T \) are diagonally dominant. Thus \( M_k^{-1}(a_k) \) is positive semi-definite.

Then \( f_k(\mu_k) \) is with each element \( f_{k,l}(\mu_k) \) given as

\[
f_{k,l}(\mu_k) = \Delta P_{k,l}^0 - e_{m_{k,l}}^T \cdot M_k^{-1}(a_k) \cdot \tilde{b}_k(\mu_k).
\]

The monotonicity condition (4.38) then can be written as

\[
0 \leq \sum_{l \in L_k} (\mu_l - \nu_l)(f_{k,l}(\mu) - f_{k,l}(\nu)) = \sum_{l \in L_k} (\mu_l - \nu_l) \cdot e_{m_{k,l}}^T \cdot M_k^{-1}(a_k) \cdot (\tilde{b}_k(\mu) - \tilde{b}_k(\nu)). \tag{4.39}
\]

By using (4.28), each element in \( \tilde{b}_k(\mu) - \tilde{b}_k(\nu) \) is

\[
\tilde{b}_{k,i}(\mu) - \tilde{b}_{k,i}(\nu) = \frac{\beta_{k,i}(\mu - \nu)}{(\alpha_{k,i} + \mu_{r(i)}) (\alpha_{k,i} + \nu_{r(i)})}, \quad \forall \ i \in M_k.
\]

We define a diagonal matrix \( diag(\tilde{\beta}_k(\mu, \nu)) \) with each diagonal element given as

\[
\tilde{\beta}_{k,i}(\mu, \nu) = \frac{\beta_{k,i}}{(\alpha_{k,i} + \mu_{r(i)}) (\alpha_{k,i} + \nu_{r(i)})} > 0, \quad \forall \ i \in M_k.
\]
Recall that we have defined a matrix $E_k$ with each column vector $e_{k,l}$ given as $e_{k,l} = e_{m_k,l}$ in (4.20), such that the $(m_k \times l_k)$ matrix $E_k$ is defined as

\[
E_{k,i,l} = \begin{cases} 
1, & \text{channel } i \text{ is added on link } l \\
0, & \text{otherwise.}
\end{cases}
\]

Hence

\[
\begin{bmatrix} 
\mu_r(1) \\
\vdots \\
\mu_r(m_k)
\end{bmatrix} = E_k \cdot \begin{bmatrix} 
\mu_1 \\
\vdots \\
\mu_k
\end{bmatrix}
\]

Thus we have a nice form of $\tilde{b}_k(\mu) - \tilde{b}_k(\nu)$:

\[
\tilde{b}_k(\mu) - \tilde{b}_k(\nu) = \text{diag}(\tilde{\beta}_k(\mu, \nu)) \cdot E_k \cdot (\mu - \nu),
\]

which, when substituted into (4.39), yields

\[
(\mu - \nu)^T \cdot E_k^T \cdot M_k^{-1}(a_k) \cdot \text{diag}(\tilde{\beta}_k(\mu, \nu)) \cdot E_k \cdot (\mu - \nu) \geq 0.
\]  

We note that $E_k$ and $E_k^T$ are matrices with elements being 0 or 1 and $\text{diag}(\tilde{\beta}_k(\mu, \nu))$ is a diagonal matrix with positive diagonal elements, given $\mu, \nu \in S_k$. We also have the result that $M_k^{-1}(a_k)$ is positive semi-definite. Therefore, $E_k^T \cdot M_k^{-1}(a_k) \cdot \text{diag}(\tilde{\beta}_k(\mu, \nu)) \cdot E_k$ is positive semi-definite, which leads to the result that the monotonicity condition (4.38) holds.

Since $S_k$ is nonempty, compact and convex and $f_k$ is continuous, the existence of a solution of $VI(S_k, f_k)$ is assured by Proposition A.3.

This type of variational inequality problems can be solved by using the extragradient method firstly proposed by Korpelevich [40], which is a modification of the projection algorithm [16]. Then at each stage $k$, after every $N_k$ iterations of the channel algorithm (4.31), the new link price $\mu_k(t+1)$ is generated at each iteration time $(t+1)$, according to the following link algorithm:

\[
\mu_k(t+1) = [\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t))]^+,
\]  

(4.41)
where \([\cdot]^+\) denotes the projection on \(S_k\) and \(\bar{\mu}_k(t)\) is a prediction of \(\mu_k(t+1)\):

\[
\bar{\mu}_k(t) = [\mu_k(t) - \eta_k f_k(\mu_k(t))]^+,
\]

and \(\eta_k > 0\) is the step-size.

Then from Proposition A.2, \(\mu_k^*\) is a solution of \(\text{VI}(S_k, f_k)\) if and only if

\[
[\mu_k^* - \eta_k f_k(\mu_k^*)]^+ = \mu_k^*.
\] (4.42)

The following theorem gives sufficient conditions for the link algorithm (4.41) to converge to a solution of \(\text{VI}(S_k, f_k)\).

**Theorem 4.3** Suppose that for all \(i \in \mathcal{M}_k\), \(a_{k,i}\) are selected such that (4.37) is satisfied. Then for any \(\eta_k \in (0, \frac{1}{C_k})\), where \(C_k\) is the Lipschitz constant of \(f_k\), the link algorithm (4.41) converges to a solution of \(\text{VI}(S_k, f_k)\).

The proof is divided into two parts. The first part shows that the conditions (4.36), (4.38) and the fact \(\mu_k^*\) is a solution of \(\text{VI}(S_k, f_k)\) imply that

\[
\|\mu_k(t+1) - \mu_k^*\|^2 \leq \|\mu_k(t) - \mu_k^*\|^2 - (1 - \eta_k^2 C_k^2)\|\mu_k(t) - \bar{\mu}_k(t)\|^2.
\] (4.43)

The proof of this part is adapted from discussions in [16]. The second part, i.e., the convergence result follows from the proof of Theorem 2.1 in [31].

**Proof:** (Part I) First, \(f_k\) is Lipschitz continuous on \(S_k\) with Lipschitz constant \(C_k\). Moreover, for \(a_{k,i}\) satisfying (4.37), Lemma 4.2 implies that the monotonicity condition (4.38) holds. Then we have

\[
(\bar{\mu}_k(t) - \mu_k^*)^T (f_k(\bar{\mu}_k(t)) - f_k(\mu_k^*)) \geq 0.
\] (4.44)

The fact that \(\mu_k^*\) is a solution of \(\text{VI}(S_k, f_k)\) implies that

\[
(\bar{\mu}_k(t) - \mu_k^*)^T f_k(\mu_k^*) \geq 0.
\] (4.45)

By using (4.45), the inequality (4.44) implies that

\[
(\bar{\mu}_k(t) - \mu_k^*)^T f_k(\bar{\mu}_k(t)) \geq 0.
\]
Then using $\bar{\mu}_k(t) - \mu_k^* = (\bar{\mu}_k(t) - \mu_k(t + 1)) - (\mu_k^* - \mu_k(t + 1))$ yields

$$
(\mu_k^* - \mu_k(t + 1))^T f_k(\bar{\mu}_k(t)) \leq (\bar{\mu}_k(t) - \mu_k(t + 1))^T f_k(\bar{\mu}_k(t))
$$

(4.46)

Since $\mu_k(t + 1)$ is the projection of $\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t))$ on $S_k$ and $\mu_k^* \in S_k$, it follows from Theorem A.6 (Projection Theorem) that

$$
\|\mu_k(t + 1) - \mu_k^*\|^2 \leq \|\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t)) - \mu_k^*\|^2 - \|\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t)) - \mu_k(t + 1)\|^2
$$

$$
= \|\mu_k(t) - \mu_k^*\|^2 - \|\mu_k(t) - \mu_k(t + 1)\|^2 + 2\eta_k(\mu_k^* - \mu_k(t + 1))^T f_k(\bar{\mu}_k(t)).
$$

(4.47)

By using (4.46) and $\mu_k(t) - \mu_k(t + 1) = (\mu_k(t) - \bar{\mu}_k(t)) + (\bar{\mu}_k(t) - \mu_k(t + 1))$, we have

$$
\|\mu_k(t + 1) - \mu_k^*\|^2 \leq \|\mu_k(t) - \mu_k^*\|^2 - \|\mu_k(t) - \bar{\mu}_k(t)\|^2 - \|\bar{\mu}_k(t) - \mu_k(t + 1)\|^2
$$

$$
+ 2(\mu_k(t + 1) - \bar{\mu}_k(t))^T (\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t)) - \bar{\mu}_k(t)).
$$

Note that

$$
\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t)) - \bar{\mu}_k(t) = (\mu_k(t) - \eta_k f_k(\mu_k(t)) - \bar{\mu}_k(t)) + \eta_k (f_k(\mu_k(t)) - f_k(\bar{\mu}_k(t))).
$$

Thus it follows that for the last term on the right hand side of (4.47), we can write

$$
(\mu_k(t + 1) - \bar{\mu}_k(t))^T (\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t)) - \bar{\mu}_k(t))
$$

$$
= (\mu_k(t + 1) - \bar{\mu}_k(t))^T (\mu_k(t) - \eta_k f_k(\mu_k(t)) - \bar{\mu}_k(t))
$$

$$
+ \eta_k (\mu_k(t + 1) - \bar{\mu}_k(t))^T (f_k(\mu_k(t)) - f_k(\bar{\mu}_k(t)))
$$

$$
\leq \eta_k (\mu_k(t + 1) - \bar{\mu}_k(t))^T (f_k(\mu_k(t)) - f_k(\bar{\mu}_k(t))),
$$

where the last inequality follows from the fact that $\bar{\mu}_k(t)$ is the projection of $\mu_k(t) - \eta_k f_k(\mu_k(t))$ on $S_k$, $\mu_k(t + 1) \in S_k$ and Theorem A.6 (Projection Theorem).

Using the Lipschitz continuity condition (4.36) yields

$$
(\mu_k(t + 1) - \bar{\mu}_k(t))^T (\mu_k(t) - \eta_k f_k(\bar{\mu}_k(t)) - \bar{\mu}_k(t))
$$

$$
\leq \eta_k C_k \|\mu_k(t + 1) - \bar{\mu}_k(t)\| \cdot \|\mu_k(t) - \bar{\mu}_k(t)\|.
$$

(4.48)
Finally substituting (4.48) into (4.47) yields
\[ \| \mu_k(t+1) - \mu_k^* \|^2 \leq \| \mu_k(t) - \mu_k^* \|^2 - \| \mu_k(t) - \bar{\mu}_k(t) \|^2 - \| \bar{\mu}_k(t) - \mu_k(t+1) \|^2 \]
\[ + 2\eta_k C_k \| \mu_k(t+1) - \bar{\mu}_k(t) \| \cdot \| \mu_k(t) - \bar{\mu}_k(t) \| \]
\[ = \| \mu_k(t) - \mu_k^* \|^2 - (1 - \eta_k^2 C_k^2) \| \mu_k(t) - \bar{\mu}_k(t) \|^2 \]
\[ - (\eta_k C_k \| \mu_k(t) - \bar{\mu}_k(t) \| - \| \bar{\mu}_k(t) - \mu_k(t+1) \|)^2 \]
\[ \leq \| \mu_k(t) - \mu_k^* \|^2 - (1 - \eta_k^2 C_k^2) \| \mu_k(t) - \bar{\mu}_k(t) \|^2. \]

Part I is proved.

(\textbf{Part II}) Firstly if \( \eta_k \in (0, \frac{1}{C_k}) \), we have \( 0 < (1 - \eta_k^2 C_k^2) < 1 \). Then from (4.43), we have
\[ \sum_{t=0}^{\infty} \| \mu_k(t) - \bar{\mu}_k(t) \|^2 \leq \frac{1}{1 - \eta_k^2 C_k^2} \| \mu_k(0) - \mu_k^* \|^2. \]
Notice that the right-hand-side of the inequality above is constant, so
\[ \lim_{t \to \infty} (\mu_k(t) - \bar{\mu}_k(t)) = 0. \]

Secondly, from (4.43), we have
\[ \| \mu_k(t+1) - \mu_k^* \|^2 \leq \| \mu_k(t) - \mu_k^* \|^2. \]
It follows that the sequence \( \{ \mu_k(t) \} \) generated by the link algorithm (4.41) is bounded. Thus let \( \tilde{\mu}_k^* \) be a cluster point of the sequence \( \{ \mu_k(t) \} \) and let the subsequence \( \{ \mu_k(t_j) \} \) converge to \( \tilde{\mu}_k^* \), i.e.,
\[ \lim_{j \to \infty} \mu_k(t_j) = \tilde{\mu}_k^*. \]
From the definition of projection, \( \mu_k(t) - \bar{\mu}_k(t) = \mu_k(t) - [\mu_k(t) - \eta_k f_k(\mu_k(t))]^+ \) is continuous. Then
\[ \lim_{j \to \infty} (\mu_k(t_j) - [\mu_k(t_j) - \eta_k f_k(\mu_k(t_j))]^+) = \tilde{\mu}_k^* - [\tilde{\mu}_k^* - \eta_k f_k(\tilde{\mu}_k^*)]^+ = 0. \]
It follows that (4.42) holds and thus \( \tilde{\mu}_k^* \) is a solution of VI(\( S_k, f_k \)).

Next we prove by contradiction that the sequence \( \{ \mu_k(t) \} \) has only one cluster point.
Assume that there is another cluster point \( \tilde{\mu}_k \) such that
\[
\delta := \| \tilde{\mu}_k - \tilde{\mu}_k^* \| \neq 0.
\]
Since \( \tilde{\mu}_k^* \) is also a cluster point of \( \{\mu_k(t)\} \), then there exists a \( t_0 > 0 \) such that
\[
\| \mu_k(t_0) - \tilde{\mu}_k^* \| \leq \frac{\delta}{2}.
\]
Moreover, \( \tilde{\mu}_k^* \) is a solution of \( \text{VI}(S_k, f_k) \). Thus it follows from (4.43) that \( \forall t \geq t_0 \),
\[
\| \mu_k(t) - \tilde{\mu}_k^* \| \leq \| \mu_k(t_0) - \tilde{\mu}_k^* \| \leq \frac{\delta}{2}.
\]
Then we have \( \forall t \geq t_0 \),
\[
\| \mu_k(t) - \tilde{\mu}_k \| \geq \| \tilde{\mu}_k - \tilde{\mu}_k^* \| - \| \mu_k(t) - \tilde{\mu}_k^* \| = \delta - \| \mu_k(t) - \tilde{\mu}_k^* \| \geq \frac{\delta}{2}
\]
This contradicts the assumption that \( \tilde{\mu}_k \) is a cluster point of \( \{\mu_k(t)\} \). Thus the sequence \( \{\mu_k(t)\} \) converges to a solution of \( \text{VI}(S_k, f_k), \tilde{\mu}_k^* \).

Remark 4.4 Substituting (4.33) into (4.41) yields
\[
\mu_k(t + 1) = [\mu_k(t) - \eta_k (\Delta P_k^0 - E_k u_k(\tilde{\mu}_k(t)))]^+,
\]
(4.49)
which implies that the link needs updated channel power when the link price is $\bar{\mu}_k(t)$ instead of $\mu_k(t)$. Thus practically the channel algorithm and link algorithm are realized as in Fig. 4.4.

**Figure 4.4: Flow of the channel and link algorithm**

**Remark 4.5** For any given $\mu_k \geq 0$, Theorem 4.2 shows that the channel algorithm converges to the NE solution which is unique in the sense of innerness. Though Theorem 4.3 does not guarantee the uniqueness of a solution, $\mu_k^*, (u_k^*, \mu_k^*)$ is an optimal NE solution-Lagrange multiplier pair and algorithm (4.41) converges to one such pair.
4.4 Simulations

In this section we present MATLAB simulation results for a multi-link topology shown in Fig. 2.6 by using the iterative hierarchical algorithm (4.31), (4.41). Each link with same number of amplified spans has an individual link capacity constraint of \( P_0^0 = 1.5 \text{ mW}, \ P_0^1 = 2.5 \text{ mW} \) and \( P_0^2 = 2.0 \text{ mW} \). Three channels are located at the following wavelengths: 1555 nm, 1558 nm, 1561 nm. All optical amplifiers have the same gain spectral shape. The gain value for each channel \( i \) on the spectral shape is given as \( G_{l,i} = 19 \text{ dB}, \ G_{l,2} = 23 \text{ dB} \) and \( G_{l,3} = 21 \text{ dB} \), where the index \( l = 1, 2, 3 \). The diagonal elements of stage system matrices are:

\[
\hat{\Gamma}_{11,1} = 0.1101 \times 10^{-3} \\
\hat{\Gamma}_{12,2} = 0.2782 \times 10^{-3} \\
\hat{\Gamma}_{13,3} = 0.0678 \times 10^{-3} \\
\hat{\Gamma}_{21,1} = 0.0442 \times 10^{-3} \\
\hat{\Gamma}_{22,2} = 0.1071 \times 10^{-3},
\]

which will be used in the channel algorithm. Initially channels 1 and 2 are added on link 1 and channel 3 is added on link 2. We partition this multi-link game into 2 stages \( (K=2) \): stage 1 is composed of links 1 and 2 and stage 2 is link 3. The dynamic adjustment parameter \( \gamma_{k,i} \) is bounded in \([\gamma_{\text{min}}, \gamma_{\text{max}}] = [0, 10] \). The channel update algorithm is repeated with \( \beta_{k,i} = 1 \) and \( a_{k,i} \) and \( \alpha_{k,i} \) are selected proportional to \( \hat{\Gamma}_{k,i,i} \):

\[
a_{k,i} = 100 \cdot \hat{\Gamma}_{k,i,i} \quad \text{and} \quad \alpha_{k,i} = 10 \cdot \hat{\Gamma}_{k,i,i},
\]

such that the condition in (4.37) is satisfied. The initial price for each link is \( \mu_{k,l} = 2 \) and the step-size is \( \eta_1 = 0.5 \) and \( \eta_2 = 0.5 \).

In the simulation, the game on stage 1 is played first. Fig. 4.5 and Fig. 4.6 show the evolution in iteration time of channel power and OSNR on stage 1, respectively.

For every \( N = 20 \) iteration, links 1 and 2 adjust their prices simultaneously via the link algorithm (4.41) and then channels readjust their powers. The evolutions in iteration time of total power and prices are shown in Fig. 4.7 and Fig. 4.8 on link 1 and link 2,
Chapter 4. Games in Multi-link Topologies

Figure 4.5: Evolution of input power on stage 1

Figure 4.6: Evolution of OSNR on stage 1
respectively.

![Evolution of total power and price on link 1, stage 1](image)

Figure 4.7: Evolution of total power and price on link 1, stage 1

After stage 1 settles down, the game on stage 2 starts to play. For every $N = 20$ iteration, link 3 adjusts its price. The evolutions are shown in Fig. 4.9, Fig. 4.10 and Fig. 4.11.

Each of these games settles down at the inner NE solution points. After that each stage determines its $\gamma^*_k$. In this case, the final value of adjustable parameters is achieved as in the following $2 \times 3$ matrix,

$$
\gamma^* = \begin{bmatrix}
0 & 0 & 0 \\
5.9303 & 0.5615 & 0
\end{bmatrix}
$$

where row indicates the stage index and column is the channel index. For channel $i$ which is added on stages directly from the transmitter, $\gamma_{k,i} = 0$, since channel power is adjusted at the transmitter. Thus $\gamma^*$ is feasible with respect to the predefined range, $[\gamma_{min}, \gamma_{max}]$. 


Figure 4.8: Evolution of total power and price on link 2, stage 1

Figure 4.9: Evolution of input power on stage 2
Figure 4.10: Evolution of OSNR on stage 2

Figure 4.11: Evolution of total power and price on stage 2
4.5 Summary

Games with coupled utility and constraints in multi-link topologies have been considered in this chapter. Theoretical computation of Nash equilibria based on Lagrangian extension and decomposition has been applied to such games. This approach has been considered in [63] for single point-to-point fiber links with the coupled link capacity constraint. Here results have been extended to games in general multi-link topologies in which constraints propagate along links. This complicates the analysis, as convexity of the propagated constraints needs to be ensured. We have showed that convexity is automatically satisfied in single-sink topologies. The general multi-link topology has been dealt with by formulating a partitioned Nash game composed of stage Nash games, where each stage has a single-sink structure. Connections between stages are realized by exploiting the flexibility of distributed optical fiber links which have extra possible adjustments at bifurcation points. An iterative hierarchical algorithm towards computing equilibria with provable convergence under certain conditions has been proposed. The remaining work is to further extend results to general network topologies such as topologies with ring structures. This will be studied in next chapter.
Chapter 5

Games in Mesh Topologies

In Chapter 4, OSNR optimization problems with coupled constraints in general multi-link topologies were studied by introducing a partitioned Nash game composed of stage Nash games. Each stage had a single-sink structure. Channel powers were adjustable at bifurcation points where channels are dropped (exit). This chapter is an extension to Chapter 4 and studies constrained OSNR optimization problems in mesh topologies.

5.1 Introduction

In optical networks, channel powers are adjustable not only at Txs but also at optical switches [2]. Based on this, in Chapter 4 we partitioned the multi-link structure into stages and each stage was with a single-sink structure. This partition did not fully exploit this power adjustment flexibility. We presented a partitioned Nash game framework to optimize channel OSNR in general multi-link topologies. The partitioned Nash game was composed of stage Nash games and each stage Nash game was played towards minimizing channel OSNR degradation.

In this chapter, we are interested in solving OSNR optimization problems in optical networks with general topologies. Recall that the partitioned Nash game defined in Chapter 4 is composed of several ladder-nested stage Nash games. The approach is
applicable to general topologies with caution.

We take the quasi-ring topology shown in Fig. 5.1 for an example. In Fig. 5.1, we use $T_{x_i}/R_{x_i}$ to indicate the Tx/Rx that channel $i$ uses. We also use $l_i$ to refer link $l$. The quasi-ring topology is a ring-type topology with partially closed loops being formed by channel optical paths [2]. In Fig. 5.1, the optical paths of channels 1 and 3 are $l_1 \rightarrow l_2 \rightarrow l_3$ and $l_3 \rightarrow l_1 \rightarrow l_2$, respectively. We notice that each of the links is the intermediate or the end on channel optical paths. Thus it is not immediate to define similar single-sink structures as in Chapter 4. In this chapter, by fully using the flexibility of power adjustment, i.e., channel powers are adjustable at each optical switch, we partition the network into stages only composed of single optical links. Another issue of concern is the interconnection among stages. By breaking the closed loop and selecting one stage as the start, stages can be placed sequentially in a ladder-nested form. One benefit of such a partition is that the convexity of coupled constraints propagated along links on each stage is automatically satisfied.

Thus we formulate a partitioned Nash game with links to solve the OSNR optimization problem in mesh topologies. In the partitioned Nash game, each link Nash game is played towards minimizing channel OSNR degradation. Based on such a partition, the hierarchical decomposition is applicable to each link Nash game, leading to a lower-
level uncoupled Nash game for channels on each link and a higher-level problem for link pricing. Computation of equilibria is based on a three-level hierarchical algorithm. Such a partition simplifies the structure of each stage, makes it regular and scalable and benefits the development of a new link pricing algorithm.

The remainder of this chapter is organized as follows. Section 5.2 provides a network model to characterize the network connection and channel optical routes. A partitioned Nash game is formulated in Section 5.3. Section 5.4 proposes a three-level iterative hierarchical algorithm to compute equilibria. Implementation on different types of network topologies is conducted and simulation results are provided in Section 5.5, followed by a summary in Section 5.6.

## 5.2 Network Model

Consider an optical network with a set \( \mathcal{L} = \{1, \ldots, L\} \) of links. A set \( \mathcal{M} = \{1, \ldots, m\} \) of channels transmitted over the network. Let \( \mathcal{M}_l = \{1, \ldots, m_l\} \) be the set of channels transmitted over link \( l \in \mathcal{L} \). Each link \( l \) is composed of \( N_l \) cascaded optical amplifiers and optical fiber spans and has \( P^0_l \) as its constant total power target. The gain for channel \( i \) on the spectral shape of amplifiers on link \( l \) is denoted by \( G_{l,i} \). The set \( \mathcal{R}_i \) denotes the route of channel \( i \) from Tx to Rx. We virtually construct a link from each Tx to the connected optical switch as a virtual optical link (VOL). Obviously the set of VOLs, \( \mathcal{L}_v \), is equivalent to \( \mathcal{M} \). Therefore, the set of all optical links and VOLs is denoted by \( \mathcal{L}' = \mathcal{L} \cup \mathcal{L}_v = \{1, \ldots, L'\} \), where \( L' = L + m \). Furthermore, two connection matrices are defined.

A channel transmission matrix is defined as \( A = [A_{l,i}]_{L' \times m} \) with

\[
A_{l,i} = \begin{cases} 
1, & \text{channel } i \text{ uses link or virtual link } l; \\
0, & \text{otherwise.}
\end{cases}
\]
A system connection matrix is defined as \( B = [B_{k,l}]_{L' \times L} \) with
\[
B_{k,l} = \begin{cases} 
1, & \text{link/virtual link } k \text{ is connected at its output to link } l, \\
i.e., \text{the channel transmission direction between } k \text{ and } l \text{ is } k \rightarrow l; \\
0, & \text{otherwise.}
\end{cases}
\]

Let \( A = [A_l] \) with \( A_l = [A_{l,1}, \ldots, A_{l,m}]^T \) and \( B = [B_l] \) with \( B_l = [B_{1,l}, \ldots, B_{L',l}]^T \). The channel transmission matrix and the system connection matrix help describe channel routing and system physical connection conditions.

Next we specifically describe an optical link extracted from the network, followed by the results characterizing the signal, noise and OSNR at the output of this link.

\[
\begin{array}{c}
p_{k,i} \\
n_{k,i}^\text{out} \\
\end{array} \xrightarrow{\gamma_{l,i}} \begin{array}{c}
u_{l,i} \\
n_{l,i}^\text{in} \\
\end{array} \xrightarrow{\Gamma_l} \begin{array}{c}
u_{l,i} \\
n_{l,i}^\text{out} \\
\end{array} \xrightarrow{p_{l,i}}
\]

**Figure 5.2: Link \( l \)**

In Fig. 5.2, \( p_{l,i} \) and \( n_{l,i}^\text{out} \) are the signal and noise power at the output of link \( l \), respectively, and \( u_{l,i} \) and \( n_{l,i}^\text{in} \) are the signal and noise power at the input of link \( l \), respectively. The matrix \( \Gamma_l \) is the link system matrix of link \( l \). Recall that channel powers can be individually adjusted in the beginning of each optical link. The adjustment parameter for channel \( i \) on link \( l \) is denoted by \( \gamma_{l,i} \) which is identically bounded within \([\gamma_{\text{min}}, \gamma_{\text{max}}]\).

Let \( \gamma_l = [\gamma_{l,1}, \ldots, \gamma_{l,m}]^T \) and \( u_l = [u_{l,1}, \ldots, u_{l,m}]^T \), where \( u_{l,i} \in \Omega_i = [u_{\text{min}}, u_{\text{max}}] \) with \( u_{\text{min}} > 0 \) is the channel signal power at the input of link \( l \). For each input channel signal power at link \( l \), we have the following results.

**Lemma 5.1** The optical signal powers at the input of link \( l \) are given as
\[
u_{l,i} = A_{l,i} \gamma_{l,i} \sum_{k \in L'} (B_{k,l} A_{k,i} p_{k,i}), \quad \forall \ i \in \mathcal{M}, \ l \in \mathcal{L}, \quad (5.1)
\]
where

\[
p_{k,i} = \begin{cases} 
    p_{k,i}, & k \in \mathcal{L} \\
    u_i, & k \in \mathcal{L}_v,
\end{cases}
\]

where \(u_i\) is the signal power of channel \(i\) at Tx. In a vector form, we have

\[
u_l = \text{diag}(\gamma_i) \cdot \text{diag}(V_{l,i}^T) \cdot p,
\]

where

\[
V_{l,i} = \begin{bmatrix} 
    A_{l,i}B_{1,1}A_{1,1} & \vdots & \vdots \\
    A_{l,i}B_{L',1}A_{L',1} & \vdots & \vdots 
\end{bmatrix}
\]

and \(p = \begin{bmatrix} 
    p_1 \\
    \vdots \\
    p_m 
\end{bmatrix}\) with \(p_i = \begin{bmatrix} 
    p_{1,i} \\
    \vdots \\
    p_{L',i} 
\end{bmatrix}\).

Proof: For channel \(i\), its signal power launched into link \(l\) is transmitted either from one of previous links of link \(l\) or one of virtual links (Txs). Let \(k' \in \mathcal{L}'\) be such a link. Then \(B_{k',l} = 1\), \(A_{k',i} = 1\) and

\[
u_{l,i} = A_{l,i}\gamma_{l,i}(B_{k',l}A_{k',i}p_{k',i}).
\]

Notice that there is one and only one \(k'\) such that both \(B_{k',l}\) and \(A_{k',i}\) are non-zero. So the above equation can be rewritten as

\[
u_{l,i} = A_{l,i}\gamma_{l,i}\sum_{k \in \mathcal{L}', k \neq k'} B_{k,l}A_{k,i} p_{k,i} = A_{l,i}\gamma_{l,i} \sum_{k \in \mathcal{L}'} B_{k,l}A_{k,i} p_{k,i}
\]

Recall that \(\mathcal{L}' = \{1, \ldots, L'\}\). Thus \(u_{l,i}\) can be expressed as

\[
u_{l,i} = \gamma_{l,i} [A_{l,i}B_{1,1}A_{1,1}, \ldots, A_{l,i}B_{L',1}A_{L',1}] 
\begin{bmatrix} 
    p_{1,i} \\
    \vdots \\
    p_{L',i} 
\end{bmatrix} = \gamma_{l,i} \cdot V_{l,i}^T \cdot p_i.
\]

In vector form, this yields (5.2) immediately.

**Remark 5.1** The matrix \(\text{diag}(V_{l,i}^T)\) is an augmented system configuration matrix, indicating not only the connections between links but also channel routing conditions. From the system control point of view, for the network system with channels transmitted over it, inputs of this system are the signal power at the input side of each link and the outputs are the output signal power of each link.
The adjustment parameters $\gamma_{l,i}$ affect both signal and noise power simultaneously. Similarly to (5.1), the input noise power $n_{l,i}^{in}$ is given as

$$n_{l,i}^{in} = A_{l,i} \gamma_{l,i} \sum_{k \in L'} (B_{k,l} A_{k,i} n_{k,i}^{out}), \ \forall \ i \in M, \ l \in L,$$

(5.4)

where

$$n_{k,i}^{out} = \begin{cases} n_{k,i}^{out}, & k \in L \\ n_{i}^{0}, & k \in L_v, \end{cases}$$

where $n_{i}^{0}$ is the noise power of channel $i$ at Tx.

At the output of link $l$ with $l \in R_i$, we denote the OSNR of channel $i$ by $OSNR_{l,i} = \frac{p_{l,i}}{n_{l,i}^{in}}$. The following lemma is a naive extension of Lemma 2.1.

**Lemma 5.2** The optical signal and noise power of channel $i$ at the output of link $l$ are given as

$$p_{l,i} = H_{l,N_{l,i}} u_{l,i}$$

(5.5)

$$n_{l,i}^{out} = H_{l,N_{l,i}} n_{l,i}^{in} + \sum_{v=1}^{N_l} ASE_{l,v,i} \frac{H_{l,v,i}}{H_{l,N_{l,i}}}$$

(5.6)

where $ASE_{l,v,i}$ is ASE noise power of channel $i$ after span $s$ on link $l$ and

$$H_{l,v,i} = \frac{P_0}{\sum_{j \in M_l} (G_{l,i})^v u_{l,j}}, \ \forall \ v = 1, \ldots, N_l.$$

The OSNR of channel $i$ at the output of link $l$ is

$$OSNR_{l,i} = \frac{u_{l,i}}{n_{l,i}^{in} + \sum_{j \in M} A_{l,j} \Gamma_{l,j} u_{l,j}}, \ \forall \ i \in M_l$$

(5.7)

where $\Gamma_l = [\Gamma_{l,i,j}]$ is the system matrix of link $l$ with

$$\Gamma_{l,i,j} = \sum_{v=1}^{N_l} (G_{l,i})^{v-1} ASE_{l,v,i} \frac{P_0}{P_0}, \ \forall \ i, j \in M_l.$$

Based on (5.7), a recursive OSNR model for the network can be obtained directly, which is similar to the one developed in Chapter 4. Consider link $l$ in Fig. 5.2. The OSNR of channel $i$ at the output of link $l$ is given as

$$\frac{1}{OSNR_{l,i}} = \frac{1}{OSNR_{l,i}} + \sum_{j \in M} A_{l,j} \Gamma_{l,j} u_{l,j}, \ i \in M_l,$$

(5.8)
where link \( l' \) is the link precedent to link \( l \) for channel \( i \). Using (5.8), we define

\[
\frac{1}{\delta Q_{l,i}} = \frac{1}{OSNR_{l,i}} - \frac{1}{OSNR_{l',i}} = \sum_{j \in M} A_{l,j} \Gamma_{l,j} \frac{u_{l,j}}{u_{l,i}}.
\] (5.9)

It follows that \( \frac{1}{\delta Q_{l,i}} \) measures the OSNR degradation of channel \( i \) from link \( l' \) to link \( l \).

Thus instead of maximization of \( OSNR_i \) from Tx to Rx, we consider minimization of each \( \frac{1}{\delta Q_{l,i}} \) between links, i.e., minimization of individual OSNR degradation. In the next section, by using the similar methodology in Chapter 4, we present a partitioned Nash game composed of stage Nash games towards solving the OSNR optimization problem. Each stage is a link and the stage Nash game is formulated to minimize the OSNR degradation.

### 5.3 Partitioned Game Formulation

In this section we propose a partitioned Nash game framework to solve OSNR optimization problems with link capacity constraints in mesh topologies.

We partition the network into stages composed of single links. By breaking closed-loops formed by channel optical paths and selecting one stage as the start, stages can be placed sequentially in a ladder-nested form. Such a partition ensures that the convexity of coupled link capacity constraints propagated along links on each stage is automatically satisfied. Moreover, it simplifies the structure of each stage and makes it regular and scalable.

By partitioning, each stage is a link structure. Next we formulate a partitioned Nash game. The partitioned Nash game is composed of link Nash games and each link Nash game is played towards minimizing channel OSNR degradation. The set of OSNR degradation minimization problems on stages is related to the OSNR maximization problem from Tx to Rx.

On each link \( l \in \mathcal{L} \), the set of channels is \( \mathcal{M}_l = \{1, \ldots, m_l\} \). Consider the link
capacity constraint on link $l$:
\begin{equation}
g_l(u_l) = \sum_{i \in M} A_{l,i} u_{l,i} - P_0^l \leq 0, \; \forall \; l \in L,
\end{equation}
where $A_{l,i} = 1$ if channel $i$ is transmitted on link $l$ and otherwise, $A_{l,i} = 0$. Let $\tilde{\Omega}_l$ denote the action space on link $l$,
\begin{equation}
\tilde{\Omega}_l = \{ u_l \in \Omega | g_l(u_l) \leq 0 \},
\end{equation}
where $\Omega$ denotes the Cartesian product of $\Omega_i$. The action set of individual channel $i \in M$ is defined as the projection of $\tilde{\Omega}_l$ on channel $i$’s direction, namely,
\begin{equation}
\hat{\Omega}_{l,i}(u_{l,-i}) = \{ \xi \in \Omega_i | g_l(u_{l,-i}, \xi) \leq 0 \},
\end{equation}
where $u_{l,-i}$ is obtained by deleting $u_{l,i}$ from vector $u_l$. It can be seen that both the action space $\tilde{\Omega}_l$ and the individual action set $\hat{\Omega}_{l,i}(u_{l,-i})$ are compact and convex.

The link Nash game is played with each channel attempting to minimize its individual cost with respect to its OSNR degradation. We specify a channel cost function $J_{l,i}$ for channel $i$ on each link $l$. Similar to the channel cost function defined in general multi-link topologies in Chapter 4, $J_{l,i}$ is defined as a difference between a pricing function $A_{l,i} \alpha_{l,i} u_{l,i}$ and a utility function $U_{l,i}$ which reflects the associated channel’s OSNR degradation, namely,
\begin{equation}
J_{l,i} = A_{l,i} \alpha_{l,i} u_{l,i} - \beta_{l,i} U_{l,i}, \; \forall \; i \in M,
\end{equation}
with
\begin{equation}
U_{l,i} = \ln \left( 1 + \frac{A_{l,i} a_{l,i}}{\sigma_{l,i} - \Gamma_{l,i}} \right),
\end{equation}
where $\Gamma_{l,i}$ is given as in (5.7), $\beta_{l,i} > 0$ indicates the strength of the channel’s desire to minimize its OSNR degradation and $a_{l,i} > 0$ is for scalability. Substituting (5.9) into (5.12) yields
\begin{equation}
J_{l,i} = A_{l,i} \alpha_{l,i} u_{l,i} - \beta_{l,i} \ln \left( 1 + \frac{A_{l,i} u_{l,i}}{X_{l,-i}} \right), \; \forall \; i \in M,
\end{equation}
where $X_{l,-i} = \sum_{j \neq i, j \in M} A_{l,j} \Gamma_{l,j} u_{l,j}$. It follows that $J_{l,i}$ is continuously differentiable in its arguments and convex in $u_{l,i}$.
We denote such a link Nash game by $GAME(\mathcal{M}, \hat{\Omega}_{l,i}, J_{l,i})$. Note that the individual cost function $J_{l,i}$ is generally defined for each channel $i \in \mathcal{M}$. If channel $i$ is not transmitted on link $l$, i.e., $i \notin \mathcal{M}_l$, then $A_{l,i} = 0$ and $J_{l,i} = 0$, which means that the decision of channel $i$, $i \notin \mathcal{M}_l$, does not affect the decisions made by other channels $j \in \mathcal{M}_l$. Thus $GAME(\mathcal{M}, \hat{\Omega}_{l,i}, J_{l,i})$ is equivalently to a reduced Nash game played among $m_l$ channels. Furthermore, the existence of an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_{l,i}, J_{l,i})$ is guaranteed by Proposition 2.8.

We exploit the partitioned Nash game. By partitioning and selecting one stage as the start, stages can be sorted sequentially with the interpretation that across all stages, $\sum_{l \in \mathcal{L}} J_{l,i}$ is related to the overall OSNR degradation for channel $i$. Solutions of all $GAME(\mathcal{M}, \hat{\Omega}_{l,i}, J_{l,i}), l \in \mathcal{L}$ are interconnected similar to general multi-link topologies studied in Chapter 4. The explanation is given as follows.

Recall that channel powers are adjustable at optical switches and $\gamma_{l,i}$ is the adjustable parameter for channel $i$ on stage $l$. The vector form is $\gamma_l = [\gamma_{l,i}]$. Let $u^*_l = [u^*_{l,i}]$ be an NE solution of $GAME(\mathcal{M}, \hat{\Omega}_{l,i}, J_{l,i})$. The corresponding signal power vector at the output of link $l$ is $p^*_l = [p^*_{l,i}]$ and the corresponding augmented output power vector defined in Lemma 5.1 is $p^*$. Note that for those channel $i \notin \mathcal{M}_l$, we randomly set values of $u^*_{l,i}$ and $p^*_{l,i}$. By using (5.2) in Lemma 5.1, optimal $\gamma_{l,i}$ can be obtained by solving the corresponding component-wise equation in

$$u^*_l = diag(\gamma^*_l) \cdot diag(V^T_{l,i}) \cdot p^*.$$

Finally, the partitioned Nash game admits a solution if each $\gamma^*_{l,i} \in [\gamma_{\text{min}}, \gamma_{\text{max}}]$.

### 5.4 Hierarchical Decomposition and Algorithm

In this section, we first use Lagrangian extension presented in Chapter 2 to obtain an NE solution. Then we propose an iterative hierarchical algorithm for computation of equilibria of the partitioned Nash game. Recall that if channel $i$ is not transmitted on
link $l$, its decision does not affect the decisions made by other channels $j \in \mathcal{M}_i$. Thus $GAME(\mathcal{M}, \hat{\Omega}_{t,i}, J_{t,i})$ is equivalently to a reduced Nash game played among $m_t$ channels. The computation of equilibria is based on this fact and the algorithm is developed for the equivalent reduced Nash game, i.e., $GAME(\mathcal{M}_t, \hat{\Omega}_{t,i}, J_{t,i})$. We use the mark “−” to indicate the associated reduced vector. For example, $\bar{u}_l$ is the reduced vector obtained by removing those elements $u_{l,i}$, $i \notin \mathcal{M}_l$, from $u_l$. Sometimes we write $\bar{u}_l$ as $\bar{u}_l = (\bar{u}_{l,-i}, u_{l,i})$ with the vector $\bar{u}_{l,-i}$ obtained by deleting the $i^{th}$ element from the reduced vector $\bar{u}_l$.

5.4.1 Hierarchical Decomposition

On each link $l \in \mathcal{L}$, $GAME(\mathcal{M}_t, \hat{\Omega}_{t,i}, J_{t,i})$ admits an NE solution from Proposition 2.8. An NE solution can be computed by using the Lagrangian extension to the game theoretical framework presented in Chapter 2. The Lagrangian extension was used in Chapter 4 to obtain an NE solution of the Nash game for multi-link single-sink topologies. Similarly here we use this method and briefly state it for completeness.

We first define several augmented functions for $GAME(\mathcal{M}_t, \hat{\Omega}_{t,i}, J_{t,i})$. The augmented NG cost function $\tilde{J}_l(\bar{u}_l; \bar{x}_l)$ is defined as

$$\tilde{J}_l(\bar{u}_l; \bar{x}_l) = \sum_{i \in \mathcal{M}_l} J_{t,i}(\bar{u}_{l,-i}, \bar{x}_{l,i}),$$

where $J_{t,i}$ is defined in (5.14) with $A_{t,i} = 1$. The two-argument constraints are given as

$$\tilde{g}_l(\bar{u}_l; \bar{x}_l) = \sum_{i \in \mathcal{M}_l} g_{l,i}(\bar{u}_{l,-i}, \bar{x}_{l,i}).$$

Both $\tilde{J}_l(\bar{u}_l; \bar{x}_l)$ and $\tilde{g}_l(\bar{u}_l; \bar{x}_l)$ are separable in the second argument $\bar{x}_l$. The associated two-argument Lagrangian function $\tilde{L}_l$ is defined as

$$\tilde{L}_l(\bar{u}_l; \bar{x}_l; \mu_l) = \tilde{J}_l(\bar{u}_l; \bar{x}_l) + \mu_l \tilde{g}_l(\bar{u}_l; \bar{x}_l),$$

(5.15)

where the scalar $\mu_l$ is the Lagrange multiplier. The associated dual cost function $D_l(\mu_l)$ is defined as

$$D_l(\mu_l) = \tilde{L}_l(\bar{u}_l^*; \bar{x}_l^*; \mu_l),$$

(5.16)
where \( u^*_l \) is such that \( u_l = u^*_l \) satisfies \( u_l = \arg \min \tilde{L}(\bar{u}_l; \bar{x}_l; \mu_l) \), where \( \bar{\Omega} \) denotes the Cartesian product of \( \Omega_i, i \in M_l \).

The \( \text{GAME}(M_l, \hat{\Omega}_{l,i}, J_{l,i}) \) is related to a constrained minimization of \( \tilde{J}_l(\bar{u}_l; \bar{x}_l) \) with respect to \( \bar{x}_l \), which admits a fixed-point solution \( u^*_l \). Individual components of \( u^*_l \) constitute an NE solution to \( \text{GAME}(M_l, \hat{\Omega}_{l,i}, J_{l,i}) \). In the Lagrangian optimality condition, \( u^*_l \) is obtained by first minimizing \( \tilde{L}_l(\bar{u}_l; \bar{x}_l; \mu_l) \) with respect to \( \bar{x}_l \). The next step involves finding a fixed-point solution \( u^*_l \) by setting \( \bar{x}_l = \bar{u}_l \). Since \( u^*_l \) depends on \( \mu_l \), an NE solution-Lagrange multiplier pair \( (\bar{u}^*_l(\mu^*_l), \mu^*_l) \) is obtained if

\[
\mu^*_l g_l(\bar{u}^*_l) = 0 \quad \text{and} \quad \mu^*_l \geq 0.
\]

The minimization of \( \tilde{L}_l(\bar{u}_l; \bar{x}_l; \mu_l) \) can be decomposed by Proposition 2.11. Since each stage \( l \) is a single optical link, which is the simplest single-sink structure, the decomposition results (Theorem 4.1) obtained in Section 4.2.3 for single-sink topologies can be directly used. Thus \( \text{GAME}(M_l, \hat{\Omega}_{l,i}, J_{l,i}) \) can be naturally decomposed into a lower-level Nash game with no coupled constraints and a higher-level link problem for pricing (Lagrange multiplier). By using Theorem 4.1, after decomposition, the lower-level Nash game is obtained with the following individual cost function,

\[
\tilde{L}_{l,i}(\bar{u}_{l,-i}, \bar{x}_{l,i}, \mu_l) = J_{l,i}(\bar{u}_{l,-i}, \bar{x}_{l,i}) + \mu_l \bar{x}_{l,i}, \quad \forall i \in M_l,
\]

where \( J_{l,i} \) is defined in (5.14) and an uncoupled action set \( \Omega_i \). We denote this game by \( \text{GAME}(M_l, \Omega_i, \tilde{L}_{l,i}) \). We notice that \( \tilde{L}_{l,i} \) (5.17) has a similar form with \( \tilde{L}_{l,i} \) (4.25) in Chapter 4. By using Corollary 4.1, we obtain the following result characterizing the NE solution of \( \text{GAME}(M_l, \Omega_i, \tilde{L}_{l,i}) \).

**Corollary 5.1** For each given \( \mu_l \geq 0 \), \( \text{GAME}(M_l, \Omega_i, \tilde{L}_{l,i}) \) admits an NE solution if \( a_{l,i} \) in \( \tilde{L}_{l,i} \) (5.17) satisfies

\[
\sum_{j \neq i, j \in M_l} \Gamma_{l,j} < a_{l,i}, \quad \forall i \in M_l.
\]
The inner NE solution $\bar{u}^*_l(\mu_l)$ is unique, given as

$$\bar{u}^*_l(\mu_l) = \tilde{\Gamma}^{-1}_l \cdot \tilde{b}_l(\mu_l)$$  \hspace{1cm} (5.19)

where $\tilde{\Gamma}_l = [\tilde{\Gamma}_{l,i,j}]$ with

$$\tilde{\Gamma}_{l,i,j} = \begin{cases} 
  a_{l,i}, & j = i, \quad i, j \in \mathcal{M}_l, \\
  \Gamma_{l,i,j}, & j \neq i
\end{cases}$$

where $\Gamma_{l,i,j}$ is defined in (5.7) and $\tilde{b}_l(\mu_l) = [b_{l,i}(\mu_l)]$ with

$$b_{l,i}(\mu_l) = \frac{a_{l,i}\beta_{l,i}}{\alpha_{l,i} + \mu_l}, \quad i \in \mathcal{M}_l.$$  

It follows from (5.19) that $\sum_{i \in \mathcal{M}_l} u^*_{l,i}(\mu_l)$ decreases when $\mu_l$ increases. This monotonicity property will be used in the development of an iterative hierarchical algorithm.

### 5.4.2 Hierarchical Algorithm

We propose an iterative hierarchical algorithm based on the NE solution (5.19) of the lower-level Nash game and coordination at the higher level.

For given $\mu_l \geq 0$, $\bar{u}^*_l(\mu_l)$ in Corollary 5.1 is a solution to solve the Lagrangian optimality condition. Recall that $(\bar{u}^*_l(\mu^*_l), \mu^*_l)$ is an NE solution-Lagrange multiplier pair if

$$\mu^*_l g_l(\bar{u}^*_l) = 0 \quad \text{and} \quad \mu^*_l \geq 0.$$  

By using (5.10), $\mu^*_l \geq 0$ can be obtained by solving

$$\mu^*_l \left( \sum_{i \in \mathcal{M}_l} u^*_{l,i}(\mu^*_l) - P^0_l \right) = 0.$$  \hspace{1cm} (5.20)

Based on these, an iterative hierarchical algorithm is proposed for both link pricing and channel power adjustment.
Chapter 5. Games in Mesh Topologies

Channel Algorithm

Based on the similarity between $GAME(\mathcal{M}_l, \Omega_i, \bar{L}_{l,i})$ and $GAME(\mathcal{M}_k, \Omega_i, \bar{L}_{k,i})$ developed for multi-link topologies in Chapter 4, we adapt the channel algorithm proposed in Chapter 4 for channel power updating in $GAME(\mathcal{M}_l, \Omega_i, \bar{L}_{l,i})$:

$$u_{l,i}(n+1) = \frac{\beta_{l,i}}{\mu_t(t)} - \left( \frac{1}{OSNR_{l,i}(n)} - \frac{1}{OSNR_{l',i}} - \Gamma_{l,i} \right) \frac{u_{l,i}(n)}{a_{l,i}}, \forall i \in \mathcal{M}_l \quad (5.21)$$

where link $l'$ is the precedent of link $l$. Note that both $OSNR_{l',i}$ and $\mu_t(t)$ are invariable during the channel iteration on link $l$. Each channel $i$ on link $l$ updates its signal power $u_{l,i}$ based on the feedback information, i.e., its OSNR at the output of link $l$, $OSNR_{l,i}$, and fixed parameters, $\mu_t(t)$ and $OSNR_{l',i}$. From Theorem 4.2, we have the following convergence result.

**Corollary 5.2** If for all $i \in \mathcal{M}_l$, $a_{l,i}$ in $\bar{L}_{l,i}$ (5.17) are selected such that (5.18) is satisfied. Then for each given $\mu_t \geq 0$, channel algorithm (5.21) converges to the inner NE solution $\bar{u}_l^*(\mu_k)$.

Link Algorithm

The link algorithm is a gradient projection algorithm $[15, 45]$, developed based on (5.20). On each link $l$, after every $N$ iterations of the channel algorithm (5.21), the new link price $\mu_l$ is generated at each iteration time $t$, according to the following link algorithm:

$$\mu_l(t + 1) = \left[ \mu_l(t) + \eta \left( \sum_{i \in \mathcal{M}_l} u_{l,i}(\mu_l(t)) - P^t_l \right) \right]^+ \quad (5.22)$$

where $\eta > 0$ is a step-size and $[z]^+ = \max\{z, 0\}$. Practically, $N$ is sufficiently large such that each channel power converges to its solution. Thus given the total power of all channels on the link $l$, the link algorithm (5.22) is completely distributed and can be implemented by individual links using only local information. Next we provide a proof for the convergence of this link algorithm. We show that the step-size $\eta$ has an explicit upper-bound for the convergence.
For notation simplicity, we let

\[ s_l(\mu_l(t)) = \sum_{j \in M_l} u_{l,j}(\mu_l(t)). \]

Sometimes we just use \( s_l(\mu_l) \) if without causing confusion. In what follows, we further omit all the subscript \( l \) in the equation. Then (5.22) becomes

\[ \mu(t + 1) = [\mu(t) + \eta (s(\mu(t)) - P^0)]^+. \tag{5.23} \]

Recall that \( s(\mu) \) is a strictly decreasing function with respect to \( \mu \). Then the derivative of \( s(\mu) \) is negative, i.e., \( s'(\mu) < 0 \). Thus there exists a unique solution \( \mu^* > 0 \) for \( s(\mu) - \Delta P^0 = 0 \), i.e., \( s(\mu^*) - P^0 = 0 \). Moreover, \( s(\mu) \) is strictly convex. It follows that \( s(\mu) > P^0 \) when \( \mu < \mu^* \), and \( s(\mu) < P^0 \) when \( \mu > \mu^* \).

We define a function of \( \mu \),

\[ \theta(\mu) := \begin{cases} 
\frac{\mu - \mu^*}{P^0 - s(\mu)} & \text{when } \mu \neq \mu^*, \\
1 - \frac{1}{-s'(\mu^*)} & \text{when } \mu = \mu^*.
\end{cases} \]

The function \( \theta(\mu) \) has the following properties.

**Lemma 5.3** The function \( \theta(\mu) \) is positive, continuous and increasing with

\[ \lim_{\mu \to 0} \theta(\mu) = 0 \text{ and } \lim_{\mu \to \infty} \theta(\mu) = \infty. \]

**Proof:** First, the function \( \theta(\mu) \) is obviously continuous everywhere except at \( \mu^* \) and

\[ \lim_{\mu \to \mu^*} \frac{\mu - \mu^*}{P^0 - s(\mu)} = \frac{1}{-s'(\mu^*)} > 0. \]

Hence, it is continuous everywhere. Second, it can be easily checked that the function is positive when \( \mu > 0 \). Moreover, since \( s(\mu) \to \infty \) as \( \mu \to 0 \) and \( s(\mu) \to 0 \) as \( \mu \to \infty \), it follows that

\[ \lim_{\mu \to 0} \frac{\mu - \mu^*}{P^0 - s(\mu)} = 0 \text{ and } \lim_{\mu \to \infty} \frac{\mu - \mu^*}{P^0 - s(\mu)} = \infty. \]
Third, taking the derivative of $\theta$ with respect to $\mu$, we obtain
\[
\theta'(\mu) = \frac{(P^0 - s(\mu)) + (\mu - \mu^*)s'(\mu)}{(P^0 - s(\mu))^2}.
\]
Recall that $s(\mu)$ is strictly convex. Thus $(P^0 - s(\mu)) + (\mu - \mu^*)s'(\mu) > 0$ and therefore $\theta'(\mu) > 0$, which implies $\theta(\mu)$ is increasing. Hence, the conclusion follows. \hfill \blacksquare

Next we define a constant $\sigma_0$ as
\[
\sigma_0 := \inf_{x \in \mathcal{I}} \frac{x - s^{-1}(2P^0 - s(x))}{P^0 - s(x)},
\]
where $\mathcal{I}$ is an interval defined as $\mathcal{I} := [s^{-1}(2P^0), \mu^*)$ with $\mu^*$ satisfying $s(\mu^*) - P^0 = 0$ and $s^{-1}$ is the inverse function of $s$. For $\mu > \mu^*$, we define a function $\sigma(\mu)$ as
\[
\sigma(\mu) := \inf_{x \in \mathcal{I} \cup [\mu^*, \mu]} \frac{x - s^{-1}(2P^0 - s(x))}{P^0 - s(x)}.
\]
It can be easily seen that $\sigma(\mu)$ exists and is positive and with $\lim_{\mu \to \mu^*} \sigma(\mu) = \sigma_0$. Moreover, $\sigma(\mu)$ is continuous and non-increasing. Hence, if $\sigma_0 > \theta(\mu^*)$, then there is a unique $\bar{\mu} > \mu^*$ such that $\sigma(\bar{\mu}) = \theta(\bar{\mu})$.

The following result proves convergence of link algorithm (5.23).

**Theorem 5.1** If $\sigma_0 \leq \theta(\mu^*)$, let $\eta < \theta(\mu^*)$; Otherwise, let $\eta < \theta(\bar{\mu})$. Then the link algorithm (5.23) converges to $\mu^*$.

**Proof:** **Case 1:** Suppose $\eta \leq \theta(\mu^*)$ under the condition $\sigma_0 \leq \theta(\mu^*)$.

Then by Lemma 5.3 we know that $\theta^{-1}(\eta) < \mu^*$ and that for $\mu \geq \theta^{-1}(\eta)$, $\eta \leq \theta(\mu)$, which implies $0 < \mu + \eta(s(\mu) - P_0) \leq \mu^*$ if $\mu \in [\theta^{-1}(\eta), \mu^*]$ and $\mu + \eta(s(\mu) - P_0) > \mu^*$ if $\mu > \mu^*$. That means for initial condition $\mu(0) \in [\theta^{-1}(\eta), \mu^*]$, it monotonically increases with upper bound $\mu^*$ and for initial condition $\mu(0) > \mu^*$, it monotonically decreases with lower bound $\mu^*$. So they converge. For initial condition $\mu(0) < \theta^{-1}(\eta)$, we obtain $\eta > \theta(\mu(0)) = \frac{\mu(0) - \mu^*}{P_0 - s(\mu(0))}$ and $\mu(0) + \eta(s(\mu(0)) - P_0) > \mu^*$ (i.e., $\mu(1) > \mu^*$). Then from above, it converges. In other words, for this initial condition, it jumps one step and then monotonically converges. Case 1 is illustrated in Fig. 5.3.
Case 2: Suppose $\theta(\mu^*) < \eta < \theta(\bar{\mu})$ under the condition $\theta(\mu^*) < \sigma_0$.

It follows from $\theta(\mu^*) < \eta$ and Lemma 5.3 that

1. for any $\mu < \mu^*$, $\eta > \theta(\mu) = \frac{\mu - \mu^*}{P_0 - s(\mu)}$ and therefore $\mu + \eta(s(\mu) - P^0) > \mu^*$, which means that with one step it jumps to above $\mu^*$;

2. for $\mu \in [\theta^{-1}(\eta), \infty)$, we have $\eta < \theta(\mu) = \frac{\mu - \mu^*}{P_0 - s(\mu)}$ and therefore $\mu + \eta(s(\mu) - P^0) > \mu^*$, which means that with one step it remains above $\mu^*$;

3. for $\mu \in (\mu^*, \theta^{-1}(\eta))$, we have $\eta > \theta(\mu) = \frac{\mu - \mu^*}{P_0 - s(\mu)}$ and therefore $\mu + \eta(s(\mu) - P^0) < \mu^*$, which means that with one step it jumps to below $\mu^*$; Furthermore, we have $\eta > \theta(\mu) = \frac{\mu - \mu^*}{P_0 - s(\mu)}$ and $\eta < \theta(\bar{\mu}) = \sigma(\bar{\mu})$, which implies by the definition, for any $\mu \in (\mu^*, \theta^{-1}(\eta))$,

$$\eta < \frac{\mu - s^{-1}(2P^0 - s(\mu))}{P^0 - s(\mu)} \implies \eta(P^0 - s(\mu)) < \mu - s^{-1}(2P^0 - s(\mu)) \implies \mu + \eta(s(\mu) - P^0) > s^{-1}(2P^0 - s(\mu)) \implies 2P^0 - s(\mu) > s(\mu + \eta(s(\mu) - P^0)) \implies P^0 - s(\mu) > s(\mu + \eta(s(\mu) - P^0) - P^0$$

In the discrete-time domain, when $\mu(t) \in (\mu^*, \theta^{-1}(\eta))$, it leads to

$$\eta|s(\mu(t)) - P^0| > \eta|s(\mu(t + 1)) - P^0|, \ \mu(t) \in (\mu^*, \theta^{-1}(\eta)).$$
Based on the above, for different initial value of $\mu$, we conclude:

(a) For $\mu(0) \in (\mu^*, \theta^{-1}(\eta))$: $\mu(2t)$ is always above $\mu^*$ and is strictly decreasing. From (2), $\mu(2t+1)$ is always below $\mu^*$. Note that $\eta < \sigma_0$. Then for $\mu \in [s^{-1}(2P^0), \mu^*)$,

$$\eta < \frac{\mu - s^{-1}(2P^0 - s(\mu))}{P^0 - s(\mu)} \implies \eta(P^0 - s(\mu)) > \mu - s^{-1}(2P^0 - s(\mu))$$

$$\implies \mu + \eta(s(\mu) - P^0) < s^{-1}(2P^0 - s(\mu))$$

$$\implies 2P^0 - s(\mu) < s(\mu + \eta(s(\mu) - P^0))$$

When $\mu < s^{-1}(2P^0)$, we have $2P^0 - s(\mu) < 0$. Then $2P^0 - s(\mu) < s(\mu + \eta(s(\mu) - P^0))$.

Hence, for $\mu < \mu^*$, we have

$$2P^0 - s(\mu) < s(\mu + \eta(s(\mu) - P^0)) \implies s(\mu) - P^0 > P^0 - s(\mu + \eta(s(\mu) - P^0)).$$

Thus when $\mu(2t) \in (\mu^*, \theta^{-1}(\eta), \eta|s(\mu(2t)) - P^0| > \eta|s(\mu(2t+1)) - P^0|$, which implies that $\mu(2t+1)$ is strictly increasing. Both subsequences $\mu(2t)$ and $\mu(2t+1)$ have to converge to $\mu^*$ by using Lemma 3.3 and the discrete-time Lyapunov stability theory [29].

(b) For $\mu(0) \geq \theta^{-1}(\eta)$: we show that it goes into the interval $[\mu^*, \theta^{-1}(\eta))$ at a finite step by contradiction. Suppose it does not go into the interval $[\mu^*, \theta^{-1}(\eta))$. That means it remains greater than $\theta^{-1}(\eta)$ infinitely. By using the link algorithm (5.23), it is strictly decreasing and will be smaller than $\theta^{-1}(\eta)$. This contradicts the assumption. Thus it goes into the interval $[\mu^*, \theta^{-1}(\eta))$. Then by (a) the sequence converges to $\mu^*$.

(c) For $\mu(0) < \mu^*$: by using (1), it jumps to above $\mu^*$ with one step. Then by (a) and/or (b) it converges to $\mu^*$.

Case 2 is illustrated in Fig. 5.4. We conclude that the link algorithm (5.23) converges to $\mu^*$ given that $\eta < \theta(\mu^*)$ if $\sigma_0 \leq \theta(\mu^*)$ or $\eta < \theta(\bar{\mu})$ if $\sigma_0 > \theta(\mu^*)$. ■
Recall that in Chapter 4, an extragradient method (4.41) has been used on each stage $k$ for the link pricing. The extragradient method is a modified projection algorithm. The sufficient conditions for its convergence depend on a Lipschitz constant which is intractable. Specifically the step-size $\eta_k$ in (4.41) is bounded by a Lipschitz constant. The link algorithm (5.23) is a simple projection algorithm. Theorem 5.1 provides an explicit upper-bound for the step-size $\eta$. The upper-bound is related to the NE solution. It is worth mentioning that the extragradient method (4.41) can be used here for every link $l$ pricing. However, the link algorithm (5.23) can not be applied for the link pricing on each stage $k$ in Chapter 4 due to the convergence issue.

5.5 Simulations

In this section, we apply the partitioned Nash game framework to study several types of network topologies and implement the hierarchical algorithm to compute the solution. The first two types of network topologies are the multi-link topology and the quasi-ring topology. Both of them are representative for selected paths extracted from a mesh configuration. The multi-link topology was studied in Chapter 4. The multi-link structure was partitioned into stages with a single-sink structure. In this section we fully use the
flexibility of channel power adjustment at each optical switch and partition the multi-link structure into stages with single links. A mesh network topology is studied next. Simulation results are given for each type of network topologies.

Some assumptions and parameters defined in simulation for all types of network topologies are presented first. As before, in all topologies, we assume that each link is with same number of amplified spans and all optical amplifiers deployed along links have the same gain spectral shape. The dynamic adjustment parameter $\gamma_{l,i}$ is bounded within $[\gamma_{\text{min}}, \gamma_{\text{max}}] = [0, 10]$ for all $l \in \mathcal{L}$ and for all $i \in \mathcal{M}$. In partitioned Nash games, individual channel cost function $J_{l,i}$ is with following parameter selection:

\[
\alpha_{l,i} = 10 \times \Gamma_{l,i}, \\
\beta_{l,i} = 1 + 0.1 \times i, \\
a_{l,i} = 50 \times l \times \Gamma_{l,i}, \quad i \in \mathcal{M}_l, \quad l \in \mathcal{L},
\]

where $\Gamma_{l,i}$ is the diagonal element in the link system matrix $\Gamma_l$. The values of $\Gamma_{l,i}$ are obtained in each individual network topology. Note that the condition (5.18) is satisfied.

### 5.5.1 Multi-link Topologies

We implement the partition approach to a simple multi-link topology with three links and three channels shown in Fig. 5.5. Initially channels 1 and 2 are added on link 1 and channel 3 is added on link 2.

Constrained OSNR optimization problems in multi-link topologies were extensively studied in Chapter 4 by introducing a partitioned Nash game with stages. For example, following the approach in Chapter 4, the network shown in Fig. 5.5 can be partitioned into 2 stages: The first stage is composed of link 1 and link 2 and the second stage is link 3. In this chapter a more natural case is considered in which channel powers are adjustable at each optical switching node. A different partition method is used in which each link is a stage. This partition simplifies the partitioned structure and the convexity.
condition is naturally satisfied, as we have discussed before.

Next we present MATLAB simulation results for the topology shown in Fig. 5.5 by applying the iterative hierarchical algorithm (5.21), (5.22). The link total power targets are $P_1^0 = 1.5 \text{ mW}$, $P_2^0 = 2.5 \text{ mW}$ and $P_3^0 = 2.0 \text{ mW}$. The diagonal elements of each link system matrix $\Gamma_l$, $l = 1, 2, 3$, are obtained as

\[
\Gamma_{1,1} = 1.139 \times 10^{-4}, \quad \Gamma_{1,2} = 3.604 \times 10^{-4}, \\
\Gamma_{2,1} = 6.804 \times 10^{-5}, \quad \Gamma_{2,2} = 2.162 \times 10^{-4}, \quad \Gamma_{2,3} = 6.839 \times 10^{-4}, \\
\Gamma_{3,1} = 8.505 \times 10^{-5}, \quad \Gamma_{3,2} = 2.703 \times 10^{-4},
\]

which will be used in the channel algorithm. The step-size in the link algorithm is $\eta = 0.8$.

After partitioning, the stage Nash game on link 1 is played first. Fig. 5.6 shows the evolution in iteration time of channel input power on link 1.

For every $N = 20$ iteration, the link adjusts its price via the link algorithm and then channels readjust their powers. The evolutions in iteration time of total power and link price are shown in Fig. 5.7. After the Nash game on link 1 settles down, the stage Nash game on link 2 starts to play and then the game on link 3. The evolutions in iteration time are shown in Fig. 5.8, Fig. 5.9, Fig. 5.10 and Fig. 5.11.
Figure 5.6: Multi-link case: Channel input power on link 1

Figure 5.7: Multi-link case: Total power and link price on link 1
Chapter 5. Games in Mesh Topologies

Figure 5.8: Multi-link case: Channel input power on link 2

Figure 5.9: Multi-link case: Total power and link price on link 2
Chapter 5. Games in Mesh Topologies

Figure 5.10: Multi-link case: Channel input power on link 3

Figure 5.11: Multi-link case: Total power and link price on link 3
The final values of adjustable parameters are achieved as

\[
\gamma^* = \begin{bmatrix}
0 & 0 & 0 \\
3.6429 & 0.6240 & 0 \\
5.1592 & 1.7449 & 0
\end{bmatrix}.
\]

For channel \(i\) which is added on link \(l\) directly from Tx, we set \(\gamma_{l,i} = 0\), since channel power is adjusted at Tx. Thus \(\gamma^*\) is feasible with respect to the predefined range \([\gamma_{\min}, \gamma_{\max}]\).

### 5.5.2 Quasi-ring Topologies

In multi-link topologies, links are inter-connected in a ladder-nested manner and the \(L\) link Nash games can be automatically played in a precedence or parallel order. Situations are different in quasi-ring topologies. We take the quasi-ring topology in Fig. 5.1 for an example. The optical paths of channels 1 and 3 are \(l_1 \rightarrow l_2 \rightarrow l_3\) and \(l_3 \rightarrow l_1 \rightarrow l_2\), respectively. Each of the links is the intermediate or the end on channel optical paths. We partition this structure into 3 stages and each stage is a link. A partitioned Nash game is formulated composed of three link Nash games. On each link Nash game, the convexity of link capacity constraints is automatically satisfied. By breaking the closed loop and selecting one link as the start, link Nash games can be played sequentially.

Take a look at the simple quasi-ring topology in Fig. 5.1(a) and three channels whose optical paths are shown in Fig. 5.1(b). We break the closed loop and select link 3 as the starting link. The unfolded configuration is shown in Fig. 5.12.

The overall recursive process is such that stage Nash games on links are played sequentially: \(l_3 \rightarrow l_1 \rightarrow l_2\). On link 3, the adjustable parameters for channels 1 and 2 are initially set as \(\gamma^*_{3,1}\) and \(\gamma^*_{3,2}\), respectively, where the superscript \(^1\) indicates the number of iteration of the game among links. The game on link 3 settles down at \(u_3^* (\mu_3^*)\) with the corresponding channel output power \(p_3^*(\mu_3^*)\). Sequentially, the game on link 1 is played with an NE solution, \(u_1^* (\mu_1^*)\). The channel output power is \(p_1^*(\mu_1^*)\). Given \(p_3^*(\mu_3^*)\), the adjustable parameter on link 1, \(\gamma^*_{1,3}\), is determined. The game on link 2 is played after
The NE solution of this game is $u^*_2(\mu^*_2)$ and the channel output power is $p^*_2(\mu^*_2)$. Then the adjustable parameters on link 2, $\gamma^*_{2,i}$, $i = 1, 2$, are determined. With the given $p^*_2(\mu^*_2)$, link 3 determines its adjustable parameters by $\gamma^*_{3,i} = \frac{u^*_{3,i}(\mu^*_3)}{p^*_{3,i}(\mu^*_2)}$, $i = 1, 2$.

Next we present MATLAB simulation results for the topology shown in Fig. 5.1. The link total power targets are $P^0_1 = 1.5$ mW, $P^0_2 = 2.5$ mW and $P^0_3 = 2.0$ mW. The diagonal elements of each link system matrix $\Gamma_l$, $l = 1, 2, 3$, are obtained as

1. $\Gamma_{1,1} = 2.166 \times 10^{-4}$, $\Gamma_{1,2} = 6.852 \times 10^{-4}$, $\Gamma_{1,3} = 2.2 \times 10^{-3}$,
2. $\Gamma_{2,1} = 3.611 \times 10^{-4}$, $\Gamma_{2,2} = 1.100 \times 10^{-3}$,
3. $\Gamma_{3,1} = 2.708 \times 10^{-4}$, $\Gamma_{3,2} = 8.565 \times 10^{-4}$, $\Gamma_{3,3} = 2.7 \times 10^{-3}$.

The step-size in the link algorithm is $\eta = 0.1$. The partitioned Nash game is played as described above. For every $N = 20$ iteration, the link adjusts its price via the link algorithm and then channels readjust their powers. Evolutions in time of channel input power, total power and link price on each link $l$ are shown in Figs. 5.13 - 5.18, respectively.

The adjustable parameters for three links and three channels are obtained as in the following $3 \times 3$ matrix,

$$
\gamma^* = \begin{bmatrix}
0 & 0 & 0.7208 \\
7.5983 & 1.4868 & 0 \\
1.5277 & 0.6042 & 0
\end{bmatrix}.
$$
The overall game settles down since \( \gamma^* \) is feasible. Note that a different starting link can be selected, say, link 1, such that games on links are played sequentially: \( l_1 \rightarrow l_2 \rightarrow l_3 \). Typically we select the starting link where channels are added directly from Txs.

![Figure 5.13: Quasi-ring case: Channel input power on link 3](image)

5.5.3 Mesh Topologies

We study a mesh network topology as shown in Fig. 5.19(a), where 8 channels are transmitted over 6 links. The channel routes are shown in Fig. 5.19(b).

It can be seen from Fig. 5.19(b) that there exists a closed-loop among links 1, 4 and 6, which is formed by the optical paths of channels 3, 4, 5 and 6. We break the closed loop and select link 1 as the starting link. The unfold configuration is shown in Fig. 5.20.

The overall recursive play process is described as follows. The games on links 1, 6 and 4 are played in a precedence order: \( l_1 \rightarrow l_6 \rightarrow l_4 \). Since the closed loop among
Figure 5.14: Quasi-ring case: Total power and link price on link 3

Figure 5.15: Quasi-ring case: Channel input power on link 1
Chapter 5. Games in Mesh Topologies

Figure 5.16: Quasi-ring case: Total power and link price on link 1

Figure 5.17: Quasi-ring case: Channel input power on link 2
links 1, 4 and 6 is unfolded, on link 1 the adjustable parameters for channels 5 and 6 are initially set as $\gamma_{1,5}$ and $\gamma_{1,6}$, respectively. The games on links 1, 2 and 5 can be played in a parallel order, while the game on link 3 is played after games on links 2 and 5 settle down.
The last game to play is the one on link 4. After all the games settle down on $u^*_l(\mu^*_l)$ with corresponding channel output powers $p^*_l(\mu^*_l)$ and adjustable parameters $\gamma^*_l$, link 1 re-determines its adjustable parameter $\gamma^*_1$ according to $p^*_4(\mu^*_4)$ because of the closed loop among links 1, 4 and 6. Again we note that link 1 is not the only choice for the starting link. Typically we select the link where channels are added directly from transmitters as the starting link.

The simulation is executed with the user-defined step-size $\eta = 0.6$. The link total power targets are $P^0 = [2.0 1.5 2.0 2.5 1.5 1.5]$ mW. The diagonal elements of each link system matrix $\Gamma_l$, $l = 1, \ldots, 6$, are obtained as $\Gamma_{1,1} = 0.0014$, $\Gamma_{1,4} = 0.0027$, $\Gamma_{1,5} = 0.0027$, $\Gamma_{1,6} = 0.0013$, $\Gamma_{2,1} = 7.212 \times 10^{-4}$, $\Gamma_{2,2} = 0.001$, $\Gamma_{3,1} = 5.409 \times 10^{-4}$, $\Gamma_{3,2} = 8.564 \times 10^{-4}$, $\Gamma_{3,7} = 8.483 \times 10^{-4}$, $\Gamma_{3,8} = 5.34 \times 10^{-4}$, $\Gamma_{4,1} = 4.327 \times 10^{-4}$, $\Gamma_{4,2} = 6.852 \times 10^{-4}$, $\Gamma_{4,3} = 0.0011$, $\Gamma_{4,4} = 0.0022$, $\Gamma_{4,5} = 0.0022$, $\Gamma_{4,6} = 0.0011$, $\Gamma_{5,7} = 0.0011$, $\Gamma_{5,8} = 7.12 \times 10^{-4}$, $\Gamma_{6,3} = 0.0018$ and $\Gamma_{6,4} = 0.0036$. We show the evolutions in iterative time of channel input power, total power and link price on links 1, 2 and 4 in Figs. 5.21 - 5.26, respectively. The final adjustable parameter values for 6 links and 8 channels are obtained as in the $6 \times 8$ matrix below,
\[ \gamma^* = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0.7067 & 1.3548 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.6893 & 0.4308 & 0 & 0 & 0 & 0 & 0.6535 & 1.0415 \\
1.1210 & 0.7271 & 0.8486 & 0.4335 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2.5839 & 1.1778 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \]

The overall game settles down since \( \gamma^* \) is feasible.

![Graph](image)

**Figure 5.21:** Mesh case: Channel input power on link 1

### 5.6 Summary

We studied the OSNR optimization problem with link capacity constraints in general topologies. The work in this chapter is an extension of Chapter 4. By fully using the
Chapter 5. Games in Mesh Topologies

Figure 5.22: Mesh case: Total power and link price on link 1

Figure 5.23: Mesh case: Channel input power on link 2
Figure 5.24: Mesh case: Total power and link price on link 2

Figure 5.25: Mesh case: Channel input power on link 4
flexibility that channel powers are adjustable at each optical switch, we formulated a partitioned Nash game composed of link Nash games to solve the OSNR optimization problem. In the partitioned Nash game, each link Nash game is played towards minimizing channel OSNR degradation. Based on such a partition, the hierarchical decomposition is applicable to each link Nash game. By selecting a starting link, the game on each link can be played sequentially and the hierarchical decomposition leads to a lower-level Nash game for channels with no coupled constraints and a higher-level problem for link pricing. Computation of equilibria is based on a three-level hierarchical algorithm. We implemented this approach in three network topologies: the multi-link topology, the quasi-ring topology and the mesh topology. Computations of equilibria based on this hierarchical algorithm were developed and simulation examples were provided.

Recall that in the partition approach proposed in Chapter 4, the multi-link structure is partitioned into stages and each stage is with a single sink structure. Compared to such
a partition, the one proposed in this chapter simplifies the structure of each stage, makes it regular and scalable and benefits the development of a new link pricing algorithm.

From Chapter 3 to Chapter 5, game-theoretic approaches have been used to solve channel OSNR optimization problems with link capacity constraints in optical networks. In a Nash game with the presence of selfishness, Nash equilibria may not optimize overall system performance. In particular, we formulated a game-theoretic framework towards optimizing channel OSNR, however, we have not taken desired channel OSNR targets into account yet. Regarding these issues, a system optimization approach will be proposed in the next chapter to solve the optimization problem with OSNR targets in point-to-point optical links. Efforts will be taken to study the efficiency of Nash equilibria numerically.
Chapter 6

Optimization with OSNR Target Constraints

In previous chapters, game-theoretic approaches were used to solve channel OSNR optimization problems with link capacity constraints in optical networks. The channel OSNR target was considered implicitly in Chapter 3 for single point-to-point WDM links. It was shown that a desired OSNR level could be obtained with a proper pricing strategy. This chapter studies constrained OSNR optimization problem from the perspective of system performance. We formulate a system optimization problem to achieve an OSNR target for each channel while satisfying the link capacity constraint. We derive conditions for the existence of a unique optimal solution, leading to a basis for an admission control scheme. We use a barrier function to relax the original constrained system problem and develop a distributed algorithm. Furthermore, the system optimization framework is used to investigate the effects of parameters in individual game cost functions in the game-theoretic framework presented in Chapter 3. This is an alternative to study the efficiency of Nash equilibria of a Nash game. We show that OSNR targets can be achieved and efficiency can be possibly improved by appropriate selection of parameters.
6.1 Introduction

The OSNR model introduced in Chapter 2 reflects that a signal can be regarded as an interfering noise for others causing OSNR degradation. Regulating channel input power at Tx aims to achieve a satisfactory channel OSNR at Rx. Such a desired OSNR, or an OSNR target of each channel is regarded as the \textit{OSNR constraint}. In Chapter 3, the OSNR optimization problem in single point-to-point fiber links was solved via a game-theoretic approach. However, the channel OSNR constraint was not considered in the framework. Alternatively, an analytical approach was provided. It has been shown that by tuning the parameters in associated game cost functions, desired OSNR targets are possibly achieved. Each selection of parameter values is regarded as a pricing mechanism.

The main goal of this chapter is to achieve a target OSNR level for each channel, while minimizing the interference and hence improving the overall performance. The link capacity constraint is also considered. Part of the results in this chapter appeared in [54]. We formulate such a problem as a system optimization problem. The system optimization framework is also used to investigate the effects of parameters in individual game cost functions in the game-theoretic framework presented in Chapter 3. It is well known that the Nash equilibria of a game may not achieve full efficiency, which is the optimal system performance. The inefficiency of the NE solution has been studied extensively [1, 22, 35, 65, 70, 73]. The resulting degree of the efficiency loss is known as the “price of anarchy” [41]. There has been an increasing literature in recent years trying to quantify the efficiency loss under separable costs [35] and non-separable costs [65], respectively. In particular, results have suggested that the selfish behavior of players in a Nash game may not degrade the system performance arbitrarily, provided a pricing mechanism is chosen properly [35, 73]. In this chapter, instead of looking from the viewpoint of degree of efficiency (“price of anarchy”), we study the efficiency by investigating the effects of parameters in individual game cost functions presented in Chapter 3. Particularly, we use the system optimization framework to measure the efficiency, which is motivated
by following facts. We show that the aggregate cost function in the game-theoretic formulation is not automatically strictly convex and the optimal solution of the associated constrained optimization problem is not immediate. We indicate that an individual cost function in the system optimization formulation has an approximate interpretation with the game cost function in the game-theoretic formulation. We compare the numerical results based on this and show the effects of pricing mechanisms.

The remainder of this chapter is organized as follows. In Section 6.2 we present a system problem and its solution. In Section 6.3 we present a barrier function to relax the original constrained system problem and propose a distributed algorithm. We provide simulation results in Section 6.4. Moreover, we study the efficiency of Nash equilibria numerically by investigating the effects of parameters in individual game cost functions presented in Chapter 3. Section 6.5 summarizes the work and contributions.

6.2 Problem Formulation

6.2.1 System Problem

We study a single point-to-point WDM link with a set of $\mathcal{M} = \{1, \ldots, m\}$ channels transmitted over the link. The signal optical power at Tx is typically bounded for each channel. That is, $u_i$ is bounded in $\Omega_i = [0, u_{\max}]$ with constat $u_{\max} > P_0$, for all $i \in \mathcal{M}$.

The system optimization problem is subject to two specific constraints: The target OSNR constraint and the link capacity constraint. Let $\hat{\gamma}_i$ be the target OSNR of channel $i$ and $\hat{\gamma} = [\hat{\gamma}_1, \ldots, \hat{\gamma}_m]^T$. The target OSNR constraint can be written as

$$OSNR_i \geq \hat{\gamma}_i, \; \forall \; i \in \mathcal{M}. \quad (6.1)$$

Another constraint is the link capacity constraint

$$\sum_{i \in \mathcal{M}} u_i - P_0 \leq 0,$$
where $P^0$ is the link total power target.

In contrast to the game-theoretic approach to the OSNR optimization problem developed in previous chapters, we consider the OSNR optimization problem from the perspective of system performance in this chapter. The problem is stated as

$$\min C(u)$$

subject to $u_i \in \Omega_i, \forall i \in \mathcal{M}$,

$$OSNR_i \geq \hat{\gamma}_i, \forall i \in \mathcal{M},$$

$$\sum_{j \in \mathcal{M}} u_j \leq P^0,$$

where $C(u)$ is the system cost function, defined as the sum of all individual costs, $C_i(u_i)$,

$$C(u) = \sum_{i \in \mathcal{M}} C_i(u_i).$$

Each individual cost function $C_i(u_i)$ is a generic cost function that satisfies the following assumption:

**Assumption 6.1** Each $C_i(u_i)$ is strictly convex, continuously differentiable and

$$\lim_{u_i \to 0} C_i(u_i) = +\infty. \quad (6.2)$$

The cost function can be defined similarly to the form of (3.7), $C_i = P_i - U_i$. The pricing function $P_i$ is a linear function of $u_i$. The utility function $U_i$ is a logarithmic function of $u_i$, which quantifies approximately the link’s demand or channel’s willingness to pay for a certain level of $OSNR_i$ based on the relationship between $OSNR_i$ and $u_i$. The relationship is illustrated approximately in Fig. 6.1.

By using the OSNR model (2.66), (6.1) can be rewritten as

$$\frac{u_i}{n_i^0 + \sum_{j \in \mathcal{M}} \Gamma_{i,j} u_j} \geq \hat{\gamma}_i,$$

or

$$u_i + \sum_{j \in \mathcal{M}} (-\hat{\gamma}_i \Gamma_{i,j}) u_j \geq n_i^0 \hat{\gamma}_i.$$
Chapter 6. Optimization with OSNR Target Constraints

The associated vector form is

\[ Tu \geq b, \quad (6.3) \]

where

\[
T = \begin{bmatrix}
1 - \tilde{\gamma}_1 \Gamma_{1,1} & -\tilde{\gamma}_1 \Gamma_{1,2} & \cdots & -\tilde{\gamma}_1 \Gamma_{1,m} \\
-\tilde{\gamma}_2 \Gamma_{2,1} & 1 - \tilde{\gamma}_2 \Gamma_{2,2} & \cdots & -\tilde{\gamma}_2 \Gamma_{2,m} \\
\vdots & \vdots & \ddots & \vdots \\
-\tilde{\gamma}_m \Gamma_{m,1} & -\tilde{\gamma}_m \Gamma_{m,2} & \cdots & 1 - \tilde{\gamma}_m \Gamma_{m,m}
\end{bmatrix}, \quad b = \begin{bmatrix}
n_1^{0} \tilde{\gamma}_1 \\
n_2^{0} \tilde{\gamma}_2 \\
\vdots \\
n_m^{0} \tilde{\gamma}_m
\end{bmatrix}.
\]

All \( T \)'s off-diagonal entries \(-\tilde{\gamma}_i \Gamma_{i,j}\) are less than zero. By Definition A.2, \( T \) is a Z-matrix.

From the total power constraint and \( u_i \geq 0, \ \forall \ i \in \mathcal{M} \), we have \( u_i \leq P^0 \). Recalling that \( u_i \) is bounded in \( \Omega_i = [0, u_{\text{max}}] \) and \( u_{\text{max}} > P^0 \), we can deduce that the conditions \( \sum_{j \in \mathcal{M}} u_j \leq P^0 \) and \( u_i \in \Omega_i, \ \forall \ i \in \mathcal{M} \) are equivalent to \( 1^T u \leq P^0 \) and \( u \geq 0 \), where \( 1 \) is the \( m \times 1 \) all ones vector. Therefore, the constraint set of the system optimization problem is \( \Omega := \{ u \in \mathbb{R}^m \mid Tu \geq b, 1^T u \leq P^0 \ \text{and} \ u \geq 0 \} \). The constrained system
optimization problem is formulated as

$$\min C(u)$$

subject to $u \in \Omega$. \hspace{1cm} (6.4)

We denote the above system optimization problem by $OPT(\Omega, C)$. The condition (6.2) in Assumption 6.1 ensures that the solution to $OPT(\Omega, C)$ does not hit $u_i = 0, \forall i \in M$.

The following result characterizes the unique solution of $OPT(\Omega, C)$.

**Theorem 6.1** If the following conditions on $\hat{\gamma}$ hold:

$$\hat{\gamma}_i < \frac{1}{\sum_{j \in M} \Gamma_{i,j}}, \ \forall \ i \in M, \hspace{1cm} (6.5)$$

where $\Gamma = [\Gamma_{i,j}]$ is the system matrix defined in (2.67) and

$$1^T \cdot \bar{T}(\hat{\gamma}) \cdot b(\hat{\gamma}) \leq P^0, \hspace{1cm} (6.6)$$

with $b(\hat{\gamma}) = [n_1^0 \hat{\gamma}_1, \ldots, n_m^0 \hat{\gamma}_m]^T$ and $\bar{T}(\hat{\gamma}) = T^{-1}(\hat{\gamma})$, then the constraint set $\Omega$ is non-empty and $OPT(\Omega, C)$ has a unique positive solution $u^{opt}$.

**Proof:** We first show that the constraint set $\Omega$ is non-empty. Note that in the link OSNR model, the system matrix $\Gamma$ (2.67) is a positive matrix, so if (6.5) is satisfied, we have

$$1 - \hat{\gamma}_i \Gamma_{i,i} > \hat{\gamma}_i \sum_{j \in M, j \neq i} \Gamma_{i,j} > 0, \ \forall \ i \in M,$$

or equivalently,

$$1 - \hat{\gamma}_i \Gamma_{i,i} > \sum_{j \in M, j \neq i} | - \hat{\gamma}_i \Gamma_{i,j}|, \ \forall \ i \in M,$$

which implies that the Z-matrix $T$ has positive diagonal entries and by Definition A.1, $T$ is strictly diagonally dominant. According to Theorem A.1, each eigenvalue of $T$ has a positive real part. Then it follows from Theorem A.3 that $T$ is an M-matrix. So it has the following properties: $T u \geq b > 0$ implies $u \geq 0$, and $T^{-1}$ is non-negative. Thus

$$u \geq T^{-1} b := \bar{T} b \hspace{1cm} (6.7)$$
and then we have $1^T u \geq 1^T \cdot \tilde{T} \cdot b$. Note that both $\tilde{T}$ and $b$ depend on $\hat{\gamma}$, i.e., $\tilde{T} = \tilde{T}(\hat{\gamma})$ and $b = b(\hat{\gamma})$. So

$$1^T \cdot u \geq 1^T \cdot \tilde{T}(\hat{\gamma}) \cdot b(\hat{\gamma}).$$

Let $\tilde{u} = \tilde{T} b$. Then $T \tilde{u} = b$. Also we have $1^T \cdot \tilde{u} = 1^T \cdot \tilde{T} b$. By (6.6), $1^T \cdot \tilde{u} \leq P^0$. It follows that

$$\tilde{u} \in \{u \in \mathbb{R}^m | T u \geq b, 1^T u \leq P^0\}.$$ 

Thus the above set is non-empty if both (6.5) and (6.6) are satisfied. Since $T u \geq b > 0$ implies $u \geq 0$, we have proved that if $\hat{\gamma}$ is selected such that (6.5) and (6.6) are satisfied, the constraint set $\Omega$ is non-empty.

Moreover, the constraint set $\Omega$ is convex and we have $0 \leq u_i \leq P_0, \forall i \in M$. So $\Omega$ is bounded. In addition, it is also closed since it consists of the intersection of half-spaces. Thus this system optimization problem is a strictly convex optimization problem on a convex compact constraint set, which following Proposition 2.2, always admits a unique globe minimum, $u^{opt}$. ■

**Example 6.1** We illustrate the constraint set $\Omega$ and the conditions (6.5) and (6.6) with a simple example in Fig. 6.2 where $m = 2$. From (6.7), we have

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \geq \begin{bmatrix} row_1(\tilde{T}) \cdot b \\ row_2(\tilde{T}) \cdot b \end{bmatrix}$$

where $row_i(A)$ is defined as the $i^{th}$ row of the matrix $A$. From Fig. 6.2, it’s ready to see that if $\sum_{i=1}^2 row_i(\tilde{T}) b \leq P^0$, the intersection point $Q$ lies in the set of the total input power constraint, i.e.,

$$(row_1(\tilde{T}) \cdot b, row_2(\tilde{T}) \cdot b) \in \{u \in \mathbb{R}^2 | u_1 + u_2 \leq P_0\}.$$ 

Therefore the constraint set $\Omega$ is non-empty. □

**Remark 6.1** Recall that $n^0$ denotes the input noise power at $Tx$ and may include external noise, such as thermal noise. If the input noise is neglected, $n^0$ includes only external
noise, which is negligible [27]. So \( b = [n_1^0 \tilde{\gamma}_1, \ldots, n_m^0 \tilde{\gamma}_m]^T \approx 0 \) and \( P_0 \geq 1^T \cdot \tilde{T} \cdot b \approx 0 \).

It means that the constraint set is non-empty under the first condition (6.5). Thus the OSNR target \( \tilde{\gamma}_i \) can be selected in a distributed way based on the first condition (6.5).

### 6.2.2 Maximum OSNR Target

Let us take a close look at the second condition (6.6). In a real network system, it is always a question how to express the conditions under certain physical constraints. Recall that \( T = I - \text{diag}(\tilde{\gamma}) \Gamma \), where \( I \) is an identity matrix. We know from the proof of Theorem 6.1 that \( T \) is an M-matrix. By Theorem A.2, \( \rho(\text{diag}(\tilde{\gamma}) \Gamma) < 1 \) and

\[
T^{-1} = (I - \text{diag}(\tilde{\gamma}) \Gamma)^{-1} = \sum_{k=0}^{\infty} \text{diag}(\tilde{\gamma}^k) \Gamma^k
\]

exists which is positive component-wise. We can rewrite (6.6) as

\[
1^T \cdot \sum_{k=0}^{\infty} \text{diag}(\tilde{\gamma}^k) \Gamma^k \cdot \text{diag}(\tilde{\gamma}) \cdot n^0 \leq P^0.
\]

(6.8)

If \( \tilde{\gamma}_i \) increases (given \( \tilde{\gamma}_j, j \neq i \)), LHS of (6.8) will increase. We can find a maximum OSNR target \( \tilde{\gamma}_{\text{max}} \) by solving the following equation:

\[
\tilde{\gamma}_{\text{max}} \cdot 1^T \cdot (I - \tilde{\gamma}_{\text{max}} \Gamma)^{-1} \cdot n^0 = P^0.
\]

(6.9)
Based on the link OSNR model, we know that the performance for each channel is interference limited. In addition, (6.9) shows that the OSNR targets significantly affect the capacity of a link: The link decides the OSNR threshold $\gamma_{\text{max}}$ by using (6.9). Any new channel with a required OSNR target no more than $\gamma_{\text{max}}$ will be admitted to transfer over the link. This idea can be used for links to develop channel admission control schemes.

6.3 Distributed Algorithm

Recall that $OPT(\Omega, C)$ is a constrained optimization problem. There are several computational methods for solving the constrained optimization problem [14]. In this section, we use a barrier (or penalty) function to relax the constrained optimization problem. Barrier functions set a barrier against leaving the feasible region: If the optimal solution occurs at the boundary of the feasible region, the procedure moves from the interior to the boundary [14]. We first show that by appropriate choice of barrier functions, the solution of the relaxed system problem can arbitrarily approximate the one of the original problem $OPT(\Omega, C)$. Then a distributed primal algorithm is presented for the relaxed system problem.

From the proof of Theorem 6.1 that $Tu \geq b > 0$ implies $u \geq 0$, $OPT(\Omega, C)$ can be rewritten succinctly as

$$\min \ C(u)$$
subject to \ $\hat{T}u \geq \hat{b}$, \quad (6.10)

where

$$\hat{T} = \begin{bmatrix} T \\ -1^T \end{bmatrix} \quad \text{and} \quad \hat{b} = \begin{bmatrix} b \\ -p^0 \end{bmatrix}.$$

6.3.1 Relaxed System Problem

A barrier function, $\lambda_i : \mathbb{R} \to \mathbb{R}$, is selected with the following properties:
(P.1) \( \forall i \in \mathcal{M}, \lambda_i(x) \) is non-increasing, continuous and

\[
\lim_{u_i \to \infty} \int_{\hat{b}_i}^{y_i(u)} \lambda_i(x)dx \to -\infty
\]  

(6.11)

where

\[
y_i(u) := \text{row}_i(\hat{T})u.
\]  

(6.12)

(P.2) \( \lambda_i(x) \) attains the value 0 if \( x > \hat{b}_i \), where \( \hat{b}_i \) is defined in (6.10). An example is illustrated in Fig. 6.3.

![Figure 6.3: A barrier function example](image)

By using a barrier function \( \lambda_i \) with properties (P.1) and (P.2), we construct a function

\[
V_p(u) = \sum_{i \in \mathcal{M}} C_i(u_i) - \sum_{i \in \mathcal{M}} \int_{\hat{b}_i}^{y_i(u)} \lambda_i(x)dx.
\]  

(6.13)

Based on \( V_p(u) \), we establish a relaxed system problem:

\[
\min_{u \geq 0} V_p(u).
\]  

(6.14)

Recall that the cost function \( C_i(u_i) \) is strictly convex and \( C(u) \) is also strictly convex. Thus the non-increasing property of the barrier function together with Assumption 6.1
ensures that \( V_p(u) \) is strictly convex [77]. Then it has a unique internal minimum value satisfying the following equations,

\[
\frac{\partial V_p(u)}{\partial u_i} = C'_i(u_i) - row_i(\hat{T}^T)\lambda(y(u)) = 0, \ \forall \ i \in \mathcal{M}.
\]

Thus solving the set of above equations we obtain the unique solution \( \bar{u}^{opt} \) of (6.14):

\[
\bar{u}_i^{opt} = C'_i^{-1}(row_i(\hat{T}^T)\lambda(y(\bar{u}^{opt}))), \ \forall \ i \in \mathcal{M}.
\]

The barrier function \( \lambda_i(\cdot) \) can be selected such that the unique solution of (6.14) may arbitrarily closely approximate the optimal solution of \( OPT(\Omega, C) \). For example, a barrier function can be defined as [38]

\[
\lambda_i(x) = \left[ \frac{\hat{b}_i - x + \epsilon}{\epsilon^2} \right]^+,
\]

where \( [x]^+ = \max\{x, 0\} \).

### 6.3.2 Primal Algorithm

Now we develop a distributed algorithm for the relaxed system problem. A primal algorithm is defined as a set of following differential equations:

\[
\dot{u}_i(t) = g_i(u_i, s_i) = -k_i \frac{\partial V_p(u)}{\partial u_i} = -k_i \left( C'_i(u_i(t)) - s_i(t) \right), \ \forall \ i \in \mathcal{M}, \quad (6.15)
\]

where the coefficient \( k_i > 0 \) and \( s_i(t) \) is defined as:

\[
s_i(t) = row_i(\hat{T}^T)\lambda(y(u(t))), \quad (6.16)
\]

where \( \lambda(\cdot) = [\lambda_1(\cdot), \ldots, \lambda_{m+1}(\cdot)]^T \) is a pre-defined barrier function vector. Each \( \lambda_i(\cdot) \) satisfies the properties (P.1) and (P.2).

The algorithm (6.15) is a gradient algorithm and can be implemented in a distributed way. Each channel varies its input power \( u_i \) gradually as in (6.15), while the link (network system) calculates the vector \( s(t) = [s_1(t), \ldots, s_m(t)]^T \) based on the received input.
powers, OSNR preference and link constraint, and then feeds this updated information back to each channel. The primal algorithm is represented in Fig. 6.4.

The following theorem states that the unique equilibrium of the algorithm (6.15) corresponds to the unique solution of (6.14), $\bar{u}^{\text{opt}}$. Moreover the solution is globally asymptotically stable.

**Theorem 6.2** The unique solution $\bar{u}^{\text{opt}}$ to the relaxed system optimization problem (6.14) is globally asymptotically stable for the system (6.15).

**Proof:** Notice that $\bar{u}^{\text{opt}}$ is the unique solution to the equations $\frac{\partial V_p(u)}{\partial u_i} = 0$, $\forall i \in \mathcal{M}$. Thus, it is the unique equilibrium point of the system (6.15).

Since $V_p(u)$ is strictly convex, it follows that $\bar{u}^{\text{opt}}$ is the global minimum point of the function $V_p(u)$. Let $C = V_p(\bar{u}^{\text{opt}})$. Then $V_p(u) > C$ for all $u \neq \bar{u}^{\text{opt}}$.

Now we construct a Lyapunov function for the system (6.15) as follows

$$V(u) = V_p(u) - C.$$ 

Then it can be easily verified that $V(u) = 0$ when $u = \bar{u}^{\text{opt}}$, and that $V(u) > 0$ when $u \neq \bar{u}^{\text{opt}}$. That is, the function $V(u)$ is positive definite with respect to the equilibrium point $u = \bar{u}^{\text{opt}}$.

Taking the derivative of $V(u)$ along the trajectory of the system gives

$$\dot{V}(u) = \sum_{i \in \mathcal{M}} \left( \frac{\partial}{\partial u_i} V_p(u) \cdot \dot{u}_i \right) = \sum_{i \in \mathcal{M}} k_i \left( \frac{\partial}{\partial u_i} V_p(u) \right)^2.$$
Thus we know that $\dot{V}(u) = 0$ when $u = \bar{u}^{opt}$, and that $\dot{V}(u) < 0$ when $u \neq \bar{u}^{opt}$. It means $V(u)$ is negative definite with respect to the equilibrium point $u = \bar{u}^{opt}$. Hence, by Lyapunov stability theory the conclusion follows.

The unique solution of the the relaxed system problem (6.14) may arbitrarily closely approximate the optimal solution of the original system problem (6.10) with an approximate selection of the barrier function.

### 6.4 Numerical Study

In this section, we first present MATLAB simulation results for a single point-to-point optical link shown in Fig. 2.5 in Chapter 2 by using the algorithm (6.15). After that, we use the system optimization framework to investigate the effects of parameters in the game-theoretic framework presented in Chapter 3. We first show the motivation is that the aggregate cost function in the game-theoretic formulation is not automatically strictly convex and the optimal solution of the associated constrained optimization problem is not immediate. We indicate that an individual cost function in the system optimization formulation has an approximate interpretation with the game cost function in the game-theoretic formulation. We then compare the numerical simulation results based on this and show the effects of parameters (pricing mechanisms).

In simulations, the link has six channels ($m = 6$) and the link total power target is $P^0 = 2.5 \text{ mW} (3.98 \text{ dBm})$. The associated system matrix $\Gamma$ is obtained as

$$
\Gamma = \begin{bmatrix}
7.438 & 7.353 & 7.269 & 7.186 & 7.103 & 6.942 \\
\end{bmatrix} \times 10^{-5}.
$$
Within the set of six channels, there are two levels of OSNR target, a 26 dB level desired on the first three channels and a 22 dB OSNR level on the next three channels. The conditions (6.5) and (6.6) on the target OSNR are satisfied. So the feasible constraint set is non-empty. The cost function for channel $i$ is defined as

$$C_i(u_i) = \alpha_i u_i - \beta_i \ln u_i,$$

(6.17)

where $\alpha_i > 0$ and $\beta_i > 0$. The selected cost function $C_i(u_i)$ is obviously strictly convex and continuously differentiable. Moreover $C_i(u_i) \to +\infty$ as $u_i \to 0$. The coefficients in (6.17) are selected as $\alpha_i = 1, i = 1, \ldots, 6$, and $\beta = [0.5, 0.51, 0.52, 0.3, 0.31, 0.32]$. Recalling the relationship between channel OSNR and channel input power shown in Fig. 6.1, the values of $\beta_i$ implicitly indicate channel OSNR preferences. The coefficient $k_i$ is fixed for each channel with $k_i = 0.01, i = 1, \ldots, 6$.

### 6.4.1 Simulation Results of System Optimization Problem

We simulate the primal algorithm (6.15). We initially set the channel power as

$$u(0) = [0.216\ 0.221\ 0.226\ 0.231\ 0.236\ 0.833] \text{ (mW)}.$$  

The barrier function is selected as

$$\lambda_i(x_i) = 1000 \left(\max\{0, \hat{b}_i - x_i\}\right)^6,$$

(6.18)

where $x_i(u) = row_i(\hat{T})u$. Notice that $\lambda_i(u_i)$ is zero when the constraints are satisfied. So there is a penalty with any violation of the constraints. The channel input power, total power and channel OSNR vs iteration time are shown in Figs. 6.5, 6.6 and 6.7, respectively. By using the primal algorithm to adjust all channel powers, two desired OSNR targets are achieved after the iterative process with the total power not exceeding the link capacity constraint.
6.4.2 Comparison of Cost Functions

Next we use the system optimization framework to investigate the effects of parameters in individual game cost functions in the game-theoretic framework presented in Chapter 3. First we show that the aggregate cost function in the game-theoretic formulation is not automatically strictly convex and the optimal solution of the associated constrained optimization problem is not immediate. Recall in Chapter 3, the cost function in \( \text{GAME}(\mathcal{M}, \mathcal{N}_i, J_i) \) is composed of a pricing function and a utility function. The utility function is defined to indicate each channel’s preference for a better OSNR. We are interested to know how much cost is added or utility is lost due to the player’s selfish behavior in a Nash game. Or in other words we consider the social welfare. There are many possible social welfare functions, one of which is the aggregate function. However,
Figure 6.6: Primal algorithm: Total power vs iteration

generally the convexity of the aggregate game cost function defined as

\[ J(u) := \sum_{i \in M} J_i(u) \]

is not longer guaranteed. The following simple example illuminates this, in which we omit the penalty term and noise in the OSNR model for simplicity.

**Example 6.2** Consider a Nash game with 3 players \((m = 3)\) with individual costs,

\[ J_i(u) = u_i - \ln \frac{u_i}{\sum_{j \neq i} u_j}, \quad i = 1, 2, 3. \]
It follows that

\[
\frac{\partial^2 J_1}{\partial u_1^2} = \frac{u_2 + u_3}{u_1} > 0
\]

\[
\frac{\partial^2 J_2}{\partial u_1^2} = -\frac{1}{(u_1 + u_3)^2} < 0
\]

\[
\frac{\partial^2 J_3}{\partial u_1^2} = -\frac{1}{(u_1 + u_2)^2} < 0
\]

Therefore,

\[
\frac{\partial^2 J}{\partial u_1^2} = \frac{u_2 + u_3}{u_1} - \frac{1}{(u_1 + u_3)^2} - \frac{1}{(u_1 + u_2)^2}
\]

The sign of \( \frac{\partial^2 J}{\partial u_1^2} \) is uncertain. Thus \( J(u) \) is not always convex with respect to \( u_1 \), even though \( J_1(u) \) is strictly convex with respect to \( u_1 \).

The constrained optimization problem associated with an aggregate cost function is not always a convex optimization problem and the optimal solutions are not immediate.
Recall that an individual cost function $C_i(u_i)$ in $OPT(\Omega, C)$ has an approximate interpretation similar to the one of the cost function $J_i(u)$ in $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. By this approximate definition, the individual cost function $C_i$ is uncoupled in $u$ for a given set of other power $u_{-i}$. Furthermore, it has an approximate interpretation similar to the one of $J_i$. Thus we build the relation between these two formulations, i.e., $OPT(\Omega, C)$ and $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ in Chapter 3. We use the central cost function in $OPT(\Omega, C)$ approximately as the welfare function of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$. We will compare simulation results based on this later. Moreover, in the next we select the system optimization framework to measure the efficiency of the NE solution numerically.

### 6.4.3 Parameter Effects in the Nash Game

Next by using the system optimization framework $OPT(\Omega, C)$, we study the effects of parameters in $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$, which is related to pricing mechanism [35, 38].

The NE solution of $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ is denoted by $u^*$. The GA developed in Chapter 3 is used, which is restated here for completeness:

$$\dot{u}_i(t) = -\mu \left( \alpha_i + \frac{1}{(P_0^0 - \sum_{j \in \mathcal{M}} u_j(t))^2} - \frac{\beta_i u_i(t)}{\text{OSNR}_i(t) + \beta_i} \right),$$

where the step-size is selected as $\mu = 0.01$.

**Remark 6.2** Theorem 3.3 states that GA converges to the NE solution if (3.56) and (3.57) are satisfied, where $u_{\min} > 0$ is a positive lower bound on each $u_i$. A lower bound on $\beta_i$ is set by (3.57). Since each channel attempts to select larger $\beta_i$ for the purpose of higher OSNR, the lower bound does not affect the results of the efficiency study.

In $OPT(\Omega, C)$, instead of using a generic $C_i(u_i)$, (6.17) is used as the cost function for channel $i$. The primal algorithm (6.15) is implemented and the barrier function is selected as in (6.18). Thus the equilibrium point of (6.15) closely approximates the solution of
$OPT(\Omega, C)$. All parameters are selected same as in Section 6.4.1. The initial channel powers are selected as

$$u(0) = [0.216 \ 0.221 \ 0.226 \ 0.231 \ 0.236 \ 0.241] \ (mW).$$

and the approximate optimal solution of $OPT(\Omega, C)$ is achieved as

$$u^{opt} = [0.5 \ 0.51 \ 0.52 \ 0.3 \ 0.31 \ 0.32] \ (mW).$$

The system cost values with respect to $u^{opt}$ is $C(u^{opt}) = 4.5789$.

We first present three cases in which the parameter selection strategy is not used as a guideline (thus it is possible that the game settles down at an NE solution where channels do not reach their OSNR targets). In all cases, the user-defined parameters $\beta_i$ in $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ are chosen as same as $\beta_i$ in $OPT(\Omega, C)$. The link sets fixed $\alpha_i$ at 0.001, 1 and 20, respectively. With these pricing mechanisms, the total power ($u_T$) vs iteration and channel OSNR vs channel number are shown in Fig. 6.8, Fig. 6.9 and Fig. 6.9 for three cases.

It is observed that without proper pricing mechanism, OSNR targets may not be achieved for some channels or all channels. The link capacity constraint is satisfied in all cases. Furthermore, we notice that the penalty term

$$\frac{1}{P^0 - \sum_{j \in \mathcal{M}} u_j}$$

in $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$ plays a key role with small $\alpha_i$. In other words, with larger $\alpha_i$ (say, $\alpha_i = 20$ in the third case), total power is smaller than the link capacity constraint. While with smaller $\alpha_i$ in the first two cases, total power approaches the constraint and higher channel OSNR is possibly achieved.

Channel powers, $u^{opt}$ and $u^*$, in three games vs channel number are shown in Fig. 6.11. The system cost $C(u) = \sum_{i \in \mathcal{M}} C_i(u_i)$ is evaluated via $u^{opt}$ and $u^*$, respectively and is shown in Table 6.1. Results imply that larger $\alpha_i$ degrades system performance and even violates the system constraints.
Figure 6.8: $u_T$ and $OSNR_i$ in game with $\alpha_i = 0.001$

Table 6.1: System cost values with different $\alpha_i$

<table>
<thead>
<tr>
<th>$\alpha_i$</th>
<th>$C(u^*)$</th>
<th>$C(u^{opt})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>4.7403</td>
<td>4.5789</td>
</tr>
<tr>
<td>1</td>
<td>4.9282</td>
<td>4.5789</td>
</tr>
<tr>
<td>20</td>
<td>9.7804</td>
<td>4.5789</td>
</tr>
</tbody>
</table>

Next we present three other cases in which proper pricing mechanisms are chosen such that OSNR targets for all channels are achieved. We compare by simulation the two approaches: system optimization approach and the game theoretical approach, proposed in this chapter and Chapter 3, respectively. In the game-theoretic framework $GAME(M, \hat{\Omega}_i, J_i)$, the parameter selection strategy (3.32) is used such that proper pricing mechanisms are chosen and OSNR targets for all channels are achieved. We present
three cases. Although the parameter selection strategy acts as a guideline for the selection of each $\beta_i$, it is practically intractable. Thus we choose proper pricing mechanism in simulation by trial and error.

The parameters $\alpha_i$ are set at 1 for all cases and $\beta_i$ are selected as in Table 6.2 such that different pricing mechanisms are chosen for $GAME(\mathcal{M}, \hat{\Omega}_i, J_i)$.

### Table 6.2: Parameters: $\beta_i$

<table>
<thead>
<tr>
<th>Game</th>
<th>$\beta_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[a]$</td>
<td>[3.8 4.8 5.8 2.6 3.0 3.5]</td>
</tr>
<tr>
<td>$[b]$</td>
<td>[5.5 7.0 9.4 4.0 4.5 5.0]</td>
</tr>
<tr>
<td>$[c]$</td>
<td>[10 12 15 8.4 8.5 8.3]</td>
</tr>
</tbody>
</table>

Since we do not use Monte Carlo method [26] to simulate, we select $\beta_i$ in three games.
by using the following rules. Firstly, $\beta_i$ increases for each channel, i.e., Game [c] has the largest $\beta_i$ compared to Game [a] and Game [b]. Secondly, Game [b] has the largest ratio of $\beta_i$ to $\beta_{\text{min}}$.

The efficiency of these two solutions $u^*$ and $u^{\text{opt}}$, is compared by evaluating the system cost $C(u)$. The corresponding system cost values are obtained and shown in Table 6.3.

<table>
<thead>
<tr>
<th>Game</th>
<th>$C(u^*)$</th>
<th>$C(u^{\text{opt}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game [a]</td>
<td>4.6171</td>
<td>4.5789</td>
</tr>
<tr>
<td>Game [b]</td>
<td>4.6216</td>
<td>4.5789</td>
</tr>
<tr>
<td>Game [c]</td>
<td>4.6057</td>
<td>4.5789</td>
</tr>
</tbody>
</table>

The results in Table 6.3 (compared with Table 6.1) verify that the efficiency in the
solution of the Nash game (user optimization) can be improved by proper pricing mechanism. The fact that no full efficiency in the solution of the Nash game is a well-known fact in the literature of economics [22], transportation [70] and network resource allocation [35]. Moreover, the Nash game solution gets very close to the optimal solution for system optimization (see Table 6.3). Furthermore, we can see that the NE solution in Game [c] is most efficient among these three cases. It implies that the efficiency can be possibly improved by appropriate selection of parameters.

Fig. 6.12 shows the total power vs iteration. Channel power and channel OSNR vs channel number are shown Fig. 6.13 and Fig. 6.14, respectively. The constraints (link capacity constraint and channel OSNR target) are satisfied in all cases. The total power in Game [c] approaches $P_0$ more than others. Moreover, we can tell from Fig. 6.14 that among the three cases, channel final OSNR values in Game [c] approach the optimal
solution of $OPT(\Omega, C)$ most.

We recall that the parameters $\beta_i$ in the Nash game are upper-bounded by the condition (3.12) in Theorem 3.1, or in other words, the ratio of $\beta_i$ to $\beta_{\text{min}}$ is upper bounded. The condition restricts each channel asking unilaterally for a much higher OSNR target than others. This phenomenon is also reflected in the selections of $\beta_i$ in three cases. We take Game [b] for an example. In this case, $\beta_3 = 9.4$ which is greatly larger than other $\beta_i$, indicating that channel 3 asks for a highest OSNR level. Meanwhile, channel 4 has the smallest $\beta_4 = 3.9$. Thus the largest ratio of $\beta_i$ to $\beta_{\text{min}}$ in Game [b] is around 2.41, which is the largest ratio among the three cases (2.23, 2.41, 1.81, in Game [a], [b], [c], respectively). Recall that Game [b] also has the largest system cost value $C(u^*)$, 4.6216. This phenomenon somewhat implies that the relative deviation of $\beta_i$ from the average may cause the degradation of system performance and less efficiency of Nash equilibrium.
In this chapter, we have studied a constrained OSNR optimization problem in optical networks from the perspective of system performance. As a first step, we have studied the single point-to-point link case. Each channel obtains a minimum OSNR level that is slightly greater than the desired one. Meanwhile, it optimizes its input power regarding target OSNR levels of all other channels and link capacity constraint. Given reasonable target OSNR levels for all channels, the system optimization problem admits a unique solution. By using a barrier function, we have relaxed the original constraint system optimization problem into an unconstrained optimization problem and a distributed primal algorithm was developed. Extension and generalization of the results from the single link case to the network case is an interesting future research direction.
Simulation results via the system optimization approach were presented. Furthermore, we used the system optimization framework to measure the efficiency of Nash equilibria of the Nash game presented in Chapter 3. We numerically investigated the effects of parameters in individual game cost functions. Simulation results have shown that OSNR target in the game-theoretic framework can be achieved and the efficiency can be possibly improved by appropriate selection of parameters. Questions are still remaining for applying Monte Carlo method numerically and an extension of this work theoretically.
Chapter 7

Experiments

Constrained OSNR optimization problems were studied theoretically and numerically in previous chapters. In particular, Chapter 3 presented a framework for a Nash game formulated towards OSNR optimization with link power capacity constraint in point-to-point fiber links. Two distributed update algorithms: PUA and GA, were developed towards finding the unique inner NE solution. Chapters 4 and 5 extended the results to network topology cases. Particularly, theoretical computation of Nash equilibria based on the Lagrangian extension was applied, which led to the development of iterative hierarchical algorithms towards computing equilibria. Consequent to theoretical results in Chapters 3 and 4, this chapter presents results of experiments on an optical network test system (ONTS) conducted to evaluate the applicability and underlying assumptions of the theoretical results.

7.1 Overview

The optical network test system (ONTS) consists of the following essential optical devices: the stabilized light source (LS), the variable optical attenuator (VOA), the optical spectrum analyzer (OSA), the optical Erbium-Doped Fiber Amplifier (OA), the tunable bandpass fiber optic filter, the fiber optical coupler and the ASE broadband source. Parts
Chapter 7. Experiments

of ONTS are shown in Fig. 7.1.

![Figure 7.1: ONTS: Devices](image)

The transmitter (Tx) is composed of a light source and a variable optical attenuator. Each channel input power is adjustable by setting the value of the corresponding VOA. An OSA provides accurate and comprehensive measurement capabilities for spectral analysis. It is used in ONTS to measure channel optical powers and OSNRs. An OA amplifies optical signals and noise simultaneously. The tunable filter is used to adjust the center wavelength of a narrow passband. It is used in ONTS to separate optical signals with different wavelengths. Fiber optical couplers are used widely in ONTS to combine or distribute optical powers from single (or multiple) inputs to single (or multiple) outputs. Couplers are designed bi-directionally and thus can be used as a coupler or a splitter. The input optical noise in ONTS is realized by using the ASE broadband source.
LabVIEW \(^1\) is used for the purpose of communication and control. GPIB (General Purpose Interface Bus, also referred to as IEEE-488) is implemented for communication. LabVIEW first communicates with light sources to initialize them (e.g., wavelength selection and light source power setting). LabVIEW communicates with VOAs to get and set the attenuation configurations and communicates with OSAs to measure the OSNR level and output power of each channel. Distributed optimization algorithms are embedded in the control block by using the MathScript Node in LabVIEW. The insertion power loss of equipments is calibrated in the control block.

### 7.2 Implementations and Results

Experiments are conducted for two cases: the single link case (shown in Fig 2.5 in Chapter 2) and the multi-link case (shown in Fig 4.2 in Chapter 4). Experiments on mesh cases currently are not conducted due to the limitation of available devices.

#### 7.2.1 Single-Link Case

The ONTS with a single optical link is setup shown in Fig 7.2.

---

\(^1\)LabVIEW (short for Laboratory Virtual Instrumentation Engineering Workbench) is a platform and development environment for a visual programming language from National Instruments®.
The link is composed of one OA. Multiple channels are transmitted with the following wavelengths: 1533.47\text{nm}, 1535.04\text{nm}, 1537.40\text{nm}, 1555.75\text{nm} and 1558.17\text{nm}. We continue using the notations defined in Chapter 3. Since only one OA is used, i.e., $N = 1$ in (2.67), diagonal elements of the system matrix $\Gamma$ are derived from (2.67), i.e.,

$$\Gamma_{i,i} = \frac{ASE_i}{P^0},$$

where $P^0 = 1.5\text{mW}$ is the constant total power target of this link and $ASE_i$ is the ASE noise, defined as

$$ASE_i = 2n_sp(G_i - 1)h\nu_iB,$$

where

- $n_sp$: amplifier excess noise factor, $n_sp = 1$
- $h$: Planck’s constant, $h = 6.626$
- $B$: optical bandwidth, $B = 10^{10}$
- $\nu_i$: optical frequency of channel $i$
- $G_i$: gain of OA at the wavelength of channel $i$

**Algorithms**

Two algorithms, PUA and GA presented in Chapter 3, are applied. In PUA, each $u_i(n+1)$ is achieved by solving the following equation:

$$\alpha_i + \frac{1}{\left(P^0 - \sum_{j \in \mathcal{M}, j \neq i} u_j(n) - u_i(n+1)\right)^2} = \frac{\beta_i a_i}{\left(1/\text{OSNR}_i(n) - \Gamma_{i,i} + a_i\right) u_i(n+1)},$$

In GA,

$$u_i(n + 1) = u_i(n) - \mu\left(\alpha_i + \frac{1}{\left(P^0 - \sum_{j \in \mathcal{M}} u_j(n)\right)^2} - \frac{\beta_i a_i}{\left(1/\text{OSNR}_i(n) + a_i - \Gamma_{i,i}\right) u_i(n)}\right),$$

where the step-size $\mu$ is selected as $\mu = 0.0025$ for $m = 2$ and $\mu = 0.0002$ for $m = 5$. We set $\alpha_i$ at 0.01 and $a_i$ at 1, respectively, for each channel. The user specific parameters $\beta_i$ are chosen as $\beta = [10, 20]$ for $m = 2$ and $\beta = [5, 10, 20, 25, 30]$ for $m = 5$. 
According to the convergence results in Chapter 3, PUA is applied to a Nash game with only two channels ($m = 2$). GA is applied to the two games with two channels ($m = 2$) and five channels ($m = 5$), respectively.

**Results**

The Nash game with two channels ($m = 2$) is conducted first. Two channels compete for the power resources via PUA and GA, respectively. Experimental data are plotted in MATLAB. The evolutions in iteration time of channel input power and total power are shown in Fig 7.3 and Fig 7.4, respectively.

![PUA on single-link: input power, m=2](image1)

**Figure 7.3: Evolution of input power: m=2**

Experiment results show that fluctuations in PUA are largely avoided in GA, which is in accordance with the theoretical and numerical results in Chapter 3. We note that there exists a difference between the results of GA and PUA in Fig 7.3, which however does not exist in simulation results. The reason is that in ONTS, when running PUA, we solve the first-order condition equation by using the measured OSNR value and input power such that the new input power is achieved. The measured data are not accurate even with calibration. However this situation does not exist when using GA. A snapshot of final results on OSA after PUA settling down is shown in Fig. 7.5.
Next the Nash game with five channels \((m = 5)\) is conducted in ONTS. The evolutions in iteration time of channel input power and total power are shown in Fig 7.6 and Fig 7.7, respectively.
Since all channels are under same conditions, for example, same noise profile, same network price $\alpha_i$, they compete for the power resource totally depending on the value of $\beta_i$: the larger value of $\beta_i$, the higher OSNR and input power values. Snapshots of initial states and final results on OSA are shown in Fig. 7.8 and Fig. 7.9, respectively.

### 7.2.2 Multi-Link Case

The ONTS with three links is setup shown in Fig 7.10. Each link is composed of an OA. The constant total power targets are $P_{10} = 1.5mW$, $P_{20} = 2.5mW$ and $P_{30} = 1.5mW$, respectively. Channels 1 and 2 are added on link 1 and channel 3 is added on link 2. We use following wavelengths for each channel: 1535.04$nm$ on Light source 1 (LS 1), 1537.40$nm$ on Light source 2 (LS 2) and 1533.47$nm$ on Light source 3 (LS 3). Channel 3 is dropped after link 2. This is realized by using filters. The filter is used to perform
wavelength selection. Channels 1 and 2 are transmitted through two filters respectively.

As proposed in Chapter 4, we partition this three-link game into two stage games ($K = 2$): Stage 1 is composed of links 1 and 2 and stage 2 is link 3. We continue using the notations defined in Chapter 4. Diagonal elements of each stage system matrix $\hat{\Gamma}_k$ are derived from stage OSNR model in Chapter 4. Particularly in this configuration,

$$\hat{\Gamma}_{1,i} = \frac{ASE_{1,i}}{P_{1}^{0}} + \frac{ASE_{2,i}}{P_{2}^{0}}, \quad i = 1, 2,$$

$$\hat{\Gamma}_{1,3} = \frac{ASE_{2,3}}{P_{2}^{0}},$$

$$\hat{\Gamma}_{2,i} = \frac{ASE_{3,i}}{P_{3}^{0}}, \quad i = 1, 2,$$
where

\[ ASE_{l,i} = 2n_{sp}(G_{l,i} - 1)\hbar\nu_i B, \]

with \( G_{l,i} \) the gain of OA at the wavelength of channel \( i \) on link \( l \).

The value of the dynamic adjustment parameter \( \gamma_{k,i} \) is in fact the attenuation value of channel \( i \) on stage \( k \), which is physically bounded.

**Algorithms**

The iterative hierarchical algorithm developed in Chapter 4 is implemented, which is composed of a channel algorithm and a link algorithm. At each stage \( k \), the following channel algorithm is used:

\[
u_{k,i}(n+1) = \frac{\beta_{k,i}}{\alpha_{k,i} + \mu_{k,i}(t)} - \left( \frac{1}{OSNR_{k,i}(n)} - \frac{1}{OSNR_{k',i}} - \hat{\Gamma}_{k,i} \right) \frac{u_{k,i}(n)}{a_{k,i}},\]
where stage $k'$ is the precedent of stage $k$ and $OSNR_{k',i}$ and $\mu_{k,r}(i)(t)$ are invariable during the channel iteration in stage $k$. Then after every $N_k$ iterations of the channel algorithm, the new link price is generated according to the following link algorithm:

$$\mu_k(t + 1) = [\mu_k(t) - \eta_k(\Delta P_k^0 - E_k u_k(\bar{\mu}_k(t)))]^+,$$
where $\Delta P^0_k$ is a modified coupled power constraint on stage $k$ and

$$\hat{\mu}_k(t) = [\mu_k(t) - \eta_k (\Delta P^0_k - E_k u_k(\mu_k(t)))^+]^+.$$ 

### Results

In the experiment, the game on Stage 1 is played first. Fig 7.11 shows the evolution in iteration time of channel input power on Stage 1.

![Figure 7.11: Evolution of input power on stage 1](image)

For every $N_1 = 10$ iteration, link 1 and link 2 adjust their prices simultaneously via the link algorithm and then channels readjust their powers. The evolutions in time of total power and prices are shown in Fig. 7.12.

After Stage 1 settles down, the game on Stage 2 starts to play. For every $N_2 = 7$ iteration, link 3 adjusts its price. The evolutions are shown in Fig. 7.13 and Fig. 7.14.
Each of these games settles down at the NE solution points.
The final values of adjustable parameters, i.e., the attenuation values, are achieved as in the following matrix,

\[
\gamma^* = \begin{bmatrix}
0 & 0 & 0 \\
4.454 & 4.073 & 0 
\end{bmatrix}.
\]

Snapshots of the values of OSNR and attenuation during the iteration on Stage 2 are shown in Fig. 7.15.

Figure 7.14: Evolution of total power and price on stage 2

Figure 7.15: Snapshot: Partial evolution of OSNR and attenuation value on stage 2
7.3 Summary

In this chapter, experiments were conducted to determine whether the theoretical results obtained in Chapters 3 and 4 could be applied practically. Two algorithms, PUA and GA, were implemented in ONTS with the single-link configuration. Partition games were performed and the iterative hierarchical algorithm was implemented in ONTS with the three-link configuration. Given the physical issues in ONTS, for example, nonlinear effects, power insertion loss and calibration, the presented results are encouraging.

Other practical issues remain for future work. Owing to the limitations of devices, we are not able to bring more than 3 links into the experiment. Channels are not distributed physically in ONTS and the distance between Tx and Rx is short, thus we have not considered the propagation time-delay issue.
Chapter 8

Summary and Conclusions

By making use of Nash game theory and other standard optimization tools, this thesis has studied constrained OSNR optimization problems in optical networks. This final chapter gives a concise summary of the thesis’ primary contributions and ends with a brief discussion about future research work.

8.1 Summary of Contributions

The goal of this thesis has been to study OSNR optimization in optical networks with constraints. Of particular interest have been the OSNR optimization problems with link capacity constraints and/or channel OSNR targets. Game theoretical approaches have been introduced to study such problems.

As a first step, we presented a game-theoretic framework for channel OSNR optimization in single point-to-point WDM links in Chapter 3. We formulated a Nash game played among channels. The link capacity constraint was considered and imposed indirectly by adding a regulation term to each cost function. Sufficient conditions were found for the existence of a unique NE solution of this game. Two iterative algorithms towards finding the NE solution in a distributed way.

Game theoretical approaches were continuously used to study OSNR optimization
problems in optical networks in Chapters 4 and 5. Games in optical networks are a subclass of network games with coupled utility and coupled constraints. The optical power propagation makes the link capacity constraints possibly non-convex from the end-to-end point of view and introduces additional complexities for analysis. The physical system constraint in optical networks provides an elimination approach to reduce constraints to a simple structure (less-coupled constraints).

The principal contribution of Chapter 4 is to propose a partition approach. The multi-link structure can be partitioned into stages at bifurcation points where channels are dropped (exit). Each stage has a single-sink structure (with either a single link or multiple links) and channel powers on each stage are adjustable. We formulated a partitioned Nash game composed of ladder-nested stage Nash games. Instead of maximization of channel OSNR from Tx to Rx (end to end), we considered minimization of channel OSNR degradation between stages. We showed that the partition directly led to the convexity of link capacity constraints on each stage. Furthermore, we applied the Lagrangian extension and decomposition results introduced in Section 2.2.2 to develop an iterative hierarchical algorithm towards computing an NE solution. More precisely, the stage Nash game was naturally decomposed into a lower-level Nash game with no coupled constraints and a higher-level problem for stage pricing. We used an extragradient method [15, 16] for stage pricing (setting prices of each link on its corresponding stage).

We looked at the mesh topology in Chapter 5 from a more natural perspective. That is, channel powers are adjustable at each optical switch. Each link is a stage after partition. The convexity of constraints propagated along links is automatically satisfied and Lagrangian extension and decomposition results can be applied directly. The other advantages are as follows.

1. By selecting starting points (links) for each quasi-ring or ring structure, it unfolds the mesh network topology with fully or partially closed loops being formed among channel optical paths.
2. It simplifies the partition structure, makes it regular and scalable. It also benefits the development of a simpler link pricing algorithm.

The link algorithm presented in Chapter 4 is a modified projection algorithm. Although it can be used in Chapter 5, we alternatively presented a gradient projection algorithm and proved its convergence. Furthermore, the proof in Chapter 5 provided an explicit upper-bound for the step-size. We showed that the upper-bound was related to the NE solution. While in Chapter 4, the step-size in the extragradient method was bounded implicitly by a Lipschitz constant.

In Chapters 3, 4 and 5, we have proposed game theoretical approaches to solve the constrained OSNR optimization problem. From the game point of view, each channel attempts to achieve a maximum OSNR or a minimum OSNR degradation. Channel OSNR targets have not been considered. Chapter 6 studied the constrained OSNR optimization problem specifically with predefined channel OSNR targets from the system point of view. More precisely, for a single point-to-point WDM link, we have considered a system OSNR optimization problem where each channel obtains its predefined OSNR target. Meanwhile, it optimizes its input power with all other channels’ OSNR targets and the link capacity constraint being taken into account. Conditions for the existence of a unique solution to the system optimization problem led to a brief discussion on a basis for an admission control scheme. We used a barrier function to relax the original system optimization problem. This led to the development of a primal algorithm in a distributed way towards computing the solution of the relaxed system problem, which approximates the optimal solution of the original system problem. This method was originally proposed in [77] and was used for power control in multicell CDMA wireless networks [6]. It was adopted for the first time to solve power control problems in optical networks. Furthermore, for a given channel OSNR target, Chapter 3 has briefly provided an idea that it is possible for channels to meet their OSNR targets by a proper pricing strategy. Chapter 6 provided a simulation comparison to study the efficiency of the NE
solution. We have verified that it is possible to improve the efficiency by tuning the parameters in cost functions.

In addition to theoretical analysis of constrained OSNR optimization problems, another focal point was the development and implementation of iterative algorithms. Extensive simulations in MATLAB have helped gain insights to the problems. Practical experimental implementations in an optical network testbed system (ONTS) have validated the applicability and underlying assumptions of the theoretical results, which is the primary contribution of Chapter 7.

8.2 Future Work

The work presented in this thesis generates several possible directions for future work.

There are couple of underlying practical issues in networks not being considered in this thesis. One key issue is time-delay, including the propagation delay of signal and the delay of feedback information. For instance, in the development of the link OSNR model (see Section 2.3.1), there exists a forward time-delay, $\tau_{i,j}$, that occurs from signal power of channel $j$, $u_j$, to OSNR of channel $i$, $OSNR_i$. Introducing the forward time-delay into the link OSNR model yields

$$OSNR_i(t) = \frac{u_i(t - \tau_{i,i})}{n_i^0 + \sum_{j \in M} \Gamma_{i,j} u_j(t - \tau_{i,j})}.$$ 

Moreover, the feedback mechanism in developing iterative algorithms makes the backward time-delay non-negligible and an example is the backward time-delay that occurs from the OSNR of channel $i$ back to its associated channel signal power. Thus the convergence of those proposed iterative algorithms is effected by time-delay. It is quite possible that the algorithms may no longer be stable given a large time-delay. It is worth mentioning that recent work by Stefanovic et al. in [78] has firstly studied the effects of time-delay in optical networks. They incorporated time-delays into the OSNR model and derived
sufficient conditions for stability of the algorithm proposed in [61]. The work in [78] can be possibly extended to solve constrained OSNR optimization problems with time-delays.

A relevant issue is *asynchronism*. Chapters 4 and 5 provided two-timescale algorithms. The channel and link algorithms converge to optimal solutions for the synchronous updating of channels’ power and links’ price, respectively. An asynchronous setting of algorithms can better resemble the reality of large-scale networks, in the sense that updates at channels and links in complex network topologies do not occur synchronously. Different asynchronous models for algorithms may lead to different convergence results. Various asynchronous algorithms with convergence properties in [16] may be possibly applied.

Another issue is related to *dynamics*. Optical networks are operated in a dynamic environment, where network reconfiguration (channels add/drop) is being performed while other existing channels are still in service. The inherent dynamics of a fiber link composed of cascaded dynamical optical elements makes the optical network dynamics an important aspect both at the data layer and the physical layer. The problem of dynamics analysis in optical networks from a system control perspective was presented in [60]. In mesh network topologies, channel routes can be changed by optical switches and closed loop structures can be formed, for example, quasi-ring structures in Chapter 5. Strictly speaking, time-delay and asynchronism are aspects of dynamics. Presented theoretical work and developed algorithms in this thesis were based on a network after reconfiguration. In other words, system parameters are stationary between any updates. From the dynamical point of view, study of the scalability of algorithms is a promising future direction.

Other possible future research ideas are as follows. Chapter 6 addresses a constrained OSNR optimization problem with OSNR targets. The problem was solved in single point-to-point fiber links. An extension to multi-link and mesh network topologies is of interest. From the end-to-end point of view, a question is arising: If the constraint set $\Omega$
in (6.4) is no longer remaining convex, is it possible to solve this constrained optimization problem via approaches based on recent non-convex study [36, 75, 81, 89]? Further to the comparison via simulation results in Chapter 6, complete theoretic analysis remains an open problem. However, the comparison via simulations in Chapter 6 offers a significant point of departure. It is also of interest to formulate the games presented in Chapters 3, 4 and 5 as a central optimization problem within the framework proposed in Chapter 6.
Appendix A

Supplementary Material

A.1 Diagonally Dominant

**Definition A.1** Let $A := [a_{ij}]$ be an $m \times m$ matrix. The matrix $A$ is said to be diagonally dominant if

$$|a_{ii}| \geq \sum_{j=1,j\neq i}^{m} |a_{ij}|, \forall i = 1, \ldots, m$$

It is said to be strictly diagonally dominant if

$$|a_{ii}| > \sum_{j=1,j\neq i}^{m} |a_{ij}|, \forall i = 1, \ldots, m$$

Some useful results are shown in the following theorem, adapted from [32].

**Theorem A.1** ([32], pp.349) Let the $m \times m$ matrix $A = [a_{ij}]$ be strictly diagonally dominant. Then $A$ is invertible and

(a) If all main diagonal entries of $A$ are positive, then all the eigenvalues of $A$ have positive real part.

(b) If $A$ is Hermitian and all main diagonal entries of $A$ are positive, then all the eigenvalues of $A$ are real and positive.
 Lemma A.1 Let $A$ be an $m \times m$ real matrix with all main diagonal entries positive. Then $A$ is positive definite if $A$ and $A^T$ are both strictly diagonally dominant.

Proof: If $A$ and $A^T$ are both strictly diagonally dominant, then it follows that

\[ a_{ii} > \sum_{j=1, j \neq i}^{m} |a_{ij}| \quad \text{and} \quad a_{ii} > \sum_{j=1, j \neq i}^{m} |a_{ji}|. \]

Thus for matrix $A_s := A + A^T$,

\[ 2a_{ii} > \sum_{j=1, j \neq i}^{m} (|a_{ij}| + |a_{ji}|) \geq \sum_{j=1, j \neq i}^{m} |a_{ij} + a_{ji}|. \]

Thus $A_s$ is also strictly diagonally dominant. From Theorem A.1, it follows that $A_s$ is positive definite. Therefore $A$ is positive definite in the sense that the symmetric part, $\frac{1}{2}A_s$, is positive definite.

A.2 M-matrix

Definition A.2 A square matrix is called a Z-matrix if all off-diagonal entries are less than or equal to zero.

Definition A.3 ( [25], Theorem 5.1) Let the $m \times m$ matrix $A = [a_{ij}]$ be a Z-matrix. $A$ is called an M-matrix if it satisfies any one of the following conditions.

1. There exists an $m \times 1$ vector $v$ with non-negative entries such that $Av > 0$.

2. Every real eigenvalue of $A$ is positive.

3. There exists a non-negative $m \times m$ matrix $B$ and a number $\lambda > \rho(B)$, where $\rho(B)$ is the spectral radius of $B$, such that $A = \lambda I - B$.

4. The real part of any eigenvalue of $A$ is positive.

5. $A$ is non-singular and the inverse of $A$ is non-negative.
6. $Av \geq 0$ implies $v \geq 0$, where $v$ is an $m \times 1$ vector.

**Theorem A.2** Let the $m \times m$ matrix $A = [a_{ij}]$ be an M-matrix. Then

1. there exists a non-negative $m \times m$ matrix $B$ and a number $\lambda > \rho(B)$, where $\rho(B)$ is the spectral radius of $B$, such that $A = \lambda I - B$.

2. specifically when $\lambda = 1$, $A = I - B$, $\rho(B) < 1$ and $(I - B)^{-1} = \sum_{k=0}^{\infty} B^k$ exists and is positive component-wise.

**Proof:** The proof for the first part can be found in [25] where Theorem 5.1 is proved. In the case when $\lambda = 1$, $\rho(B) < 1$ leads to the convergence of $\sum_{k=0}^{\infty} B^k$. Furthermore,

$$
\sum_{k=0}^{\infty} B^k(I - B) = (I + B + B^2 + \cdots)(I - B) = I
$$

Thus it follows that $(I - B)^{-1} = \sum_{k=0}^{\infty} B^k$. ■

**A.3 Maximum Theorem**

**Definition A.4** Let $S$ be a (nonempty) subset of $\mathbb{R}^n$. The set of all nonempty subsets of $S$ is called the power set of $S$, denoted by $\mathcal{P}(S)$.

**Definition A.5** Let $\Theta$ and $S$ be subsets of $\mathbb{R}^l$ and $\mathbb{R}^n$, respectively. A correspondence $\Phi$ from $\Theta$ to $S$ is a map that associates with each element $\theta \in \Theta$ a (nonempty) subset $\Phi(\theta) \subset S$.

**Definition A.6** Let $\Phi$ be a correspondence from $\Theta$ to $S$ by $\Phi : \Theta \to \mathcal{P}(S)$. $\Phi$ is said to be upper-semi-continuous at a point $\theta \in \Theta$ if for all open sets $\mathcal{V}$ such that $\Phi(\theta) \subset \mathcal{V}$, there exists an open set $\mathcal{U}$ containing $\theta$, such that $\theta' \in \mathcal{U} \cap \Theta$ implies $\Phi(\theta') \subset \mathcal{V}$. $\Phi$ is upper-semi-continuous if it is upper-semi-continuous at any point of $\Theta$. 
Example A.1  The correspondence $\Phi$ is defined by

$$\Phi(x) := \begin{cases} 
0, & \text{if } x \neq 0 \\
[-1, +1], & \text{if } x = 0 
\end{cases}$$

is upper-semi-continuous at 0.

![Figure A.1: Example: upper-semi-continuous correspondence](image)

Remark A.1  Every upper-semi-continuous single valued correspondence is a continuous function.

Theorem A.3 ( [80], Berge’s Maximum Theorem)  Let $f : X \times Y \to \mathbb{R}$ be a continuous function and $\Phi : X \to Y$ be a nonempty, compact-valued, continuous correspondence. Then

1. $f^* : X \to \mathbb{R}$ with

$$f^*(x) := \max \{ f(x, y) : y \in \Phi(x) \}$$

is a continuous function;

2. $\Phi^* : X \to Y$ with

$$\Phi^* := \arg \max \{ f(x, y) : y \in \Phi(x) \} = \{ y \in \Phi(x) : f(x, y) = f^*(x) \}$$

is a compact-valued, upper-semi-continuous correspondence.
A.4 Fixed-point Theorems

We introduce two fixed-point theorems that have been used in Chapter 2.

**Theorem A.4** ([12], Brouwer Fixed-point Theorem) *If* $S$ *is a compact and convex subset of* $\mathbb{R}^n$ *and* $f$ *is a continuous function mapping* $S$ *into itself, then there exists at least one* $x \in S$ *such that* $f(x) = x$.

**Theorem A.5** ([12], Kakutani Fixed-point Theorem) *Let* $S$ *be a compact and convex subset of* $\mathbb{R}^n$, *and let* $f$ *be an upper-semi-continuous correspondence which assigns to each* $x \in S$ *a closed and convex subset of* $S$. *Then there exists some* $x \in S$ *such that* $x \in f(x)$.

A.5 Projection Theorem

Projection method is widely used in developing iterative algorithms solving constrained optimization problems. Each time when an update jumps outside the feasible set $\mathcal{X}$, the algorithm can project it back to the set $\mathcal{X}$.

The projection of a vector $x$ onto a nonempty, closed and convex set $\mathcal{X}$ is defined with respect to the Euclidean norm, denoted by $[\cdot]^+$, i.e.,

$$[x]^+ := \arg \min_{z \in \mathcal{X}} \|z - x\|.$$

The properties of the projection are presented in the following theorem.

**Theorem A.6** ([16], Projection Theorem) *Let* $\mathcal{X}$ *be a nonempty, closed and convex subset of* $\mathbb{R}^n$.

1. *For every* $x \in \mathbb{R}^n$, *there exists a unique* $z \in \mathcal{X}$ *that minimizes* $\|z - x\|$ *over all* $z \in \mathcal{X}$ *and is denoted by* $[x]^+$.
2. Given some \( x \in \mathbb{R}^n \), a vector \( x^* \in X \) is equal to \( [x]^+ \) if and only if
\[
(z - x^*)^T(x - x^*) \leq 0, \quad \forall z \in X.
\]

3. The mapping \( f : \mathbb{R}^n \rightarrow X \) defined by \( f(x) = [x]^+ \) is continuous and nonexpansive, that is,
\[
\|[x]^+ - [y]^+\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^n.
\]

## A.6 Lipschitz Continuity

Lipschitz continuity is a smoothness condition for functions which is stronger than regular continuity. For a function \( f(x) : D \rightarrow \mathbb{R}^n \), where \( D \) is a compact subset of \( \mathbb{R}^m \), the Lipschitz condition is defined as
\[
\|f(x) - f(y)\| \leq L\|x - y\|, \quad (A.1)
\]
where \( \|\cdot\| \) denotes Euclidean norm. The positive constant \( L \) is called a Lipschitz constant.

**Definition A.7** A function \( f(x) \) is said to be Lipschitz continuous if there exists a constant \( L \geq 0 \) such that for all \( x, y \in D \), the Lipschitz condition (A.1) is satisfied.

The function \( f(x) \) is called locally Lipschitz continuous if each point of \( D \) has a neighborhood \( D_0 \) such that \( f(x) \) satisfies the Lipschitz condition for all points in \( D_0 \) with some Lipschitz constant \( L_0 \).

A locally Lipschitz continuous function \( f(x) \) on \( D \) is Lipschitz continuous on every compact (closed and bounded) set \( W \subseteq D \).

The Lipschitz property is weaker than continuous differentiability, as stated in the next proposition.

**Proposition A.1 (Lemma 3.2, [39])** If a function \( f(x) \) and \( \frac{\partial f}{\partial x}(x) \) are continuous on \( D \), then \( f(x) \) is locally Lipschitz continuous on \( D \).
A.7 Variational Inequalities

Variational inequality is a mathematical theory intended for the study of equilibrium problems. Given a subset \( X \subset \mathbb{R}^n \) and a function \( f : \mathbb{R}^n \to \mathbb{R}^n \), the variational inequality problem associated with \( X \) is to find a vector \( x^* \in X \) such that

\[
(x - x^*)^T f(x^*) \geq 0, \quad \forall \, x \in X.
\]

A short notation of this problem is denoted by \( \text{VI}(X, f) \).

A necessary and sufficient condition for a vector \( x^* \) to be a solution of \( \text{VI}(X, f) \) is given below.

**Proposition A.2** (Proposition 5.1, [16], pp.267) Let \( \gamma \) be a positive scalar and let \( G \) be a symmetric positive definite matrix. A vector \( x^* \) is a solution of \( \text{VI}(X, f) \) if and only if

\[
[x^* - \gamma G^{-1} f(x^*)]_G^+ = x^*,
\]

where \([ \cdot ]_G^+\) is the projection on \( X \) with respect to norm \( \|x\|_G = (x^T G x)^{1/2} \).

The existence of a solution is assured by the following proposition.

**Proposition A.3** (Proposition 5.2, [16], pp.268) Suppose that \( X \) is compact and that \( f : \mathbb{R}^n \to \mathbb{R}^n \) is continuous. Then there exists a solution to \( \text{VI}(X, f) \).
Appendix B

Notations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>Automatic power control</td>
</tr>
<tr>
<td>ASE</td>
<td>Amplified spontaneous emission noise</td>
</tr>
<tr>
<td>BER</td>
<td>Bit-error rate</td>
</tr>
<tr>
<td>DSL</td>
<td>Digital subscriber line</td>
</tr>
<tr>
<td>GA</td>
<td>Gradient algorithm</td>
</tr>
<tr>
<td>LHS</td>
<td>Left-hand side</td>
</tr>
<tr>
<td>LS</td>
<td>Light source</td>
</tr>
<tr>
<td>NE</td>
<td>Nash equilibrium</td>
</tr>
<tr>
<td>OA</td>
<td>Optical amplifier</td>
</tr>
<tr>
<td>OADM</td>
<td>Optical add/drop multiplexer</td>
</tr>
<tr>
<td>ONTS</td>
<td>Optical network test system</td>
</tr>
<tr>
<td>OSA</td>
<td>Optical spectrum analyzer</td>
</tr>
<tr>
<td>OSC</td>
<td>Optical service channel</td>
</tr>
</tbody>
</table>
### Appendix B. Notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSNR</td>
<td>Optical signal-to-noise ratio</td>
</tr>
<tr>
<td>OXC</td>
<td>Optical cross-connect</td>
</tr>
<tr>
<td>PUA</td>
<td>Parallel update algorithm</td>
</tr>
<tr>
<td>RHS</td>
<td>Right-hand side</td>
</tr>
<tr>
<td>Rx</td>
<td>Receiver</td>
</tr>
<tr>
<td>SIR</td>
<td>Signal to interference ratio</td>
</tr>
<tr>
<td>Tx</td>
<td>Transmitter</td>
</tr>
<tr>
<td>VOA</td>
<td>Variable optical attenuator</td>
</tr>
<tr>
<td>VOL</td>
<td>Virtual optical link</td>
</tr>
</tbody>
</table>

### Notations for signals and noises

<table>
<thead>
<tr>
<th>Location</th>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmitter</td>
<td>$u_i$</td>
<td>signal power of channel $i$</td>
</tr>
<tr>
<td></td>
<td>$n_{i}^0$</td>
<td>noise power of channel $i$</td>
</tr>
<tr>
<td>Receiver</td>
<td>$p_i$</td>
<td>signal power of channel $i$</td>
</tr>
<tr>
<td></td>
<td>$n_i$</td>
<td>noise power of channel $i$</td>
</tr>
<tr>
<td>Stage $k$</td>
<td>$u_{k,i}$</td>
<td>signal power of channel $i$ at the input</td>
</tr>
<tr>
<td></td>
<td>$n_{k,i}^{in}$</td>
<td>noise power of channel $i$ at the input</td>
</tr>
<tr>
<td></td>
<td>$p_{k,i}$</td>
<td>signal power of channel $i$ at the output</td>
</tr>
<tr>
<td></td>
<td>$n_{k,i}^{out}$</td>
<td>noise power of channel $i$ at the output</td>
</tr>
<tr>
<td>Link $l$</td>
<td>$u_{l,i}$</td>
<td>signal power of channel $i$ at the input</td>
</tr>
<tr>
<td></td>
<td>$n_{l,i}^{in}$</td>
<td>noise power of channel $i$ at the input</td>
</tr>
<tr>
<td></td>
<td>$p_{l,i}$</td>
<td>signal power of channel $i$ at the output</td>
</tr>
<tr>
<td></td>
<td>$n_{l,i}^{out}$</td>
<td>noise power of channel $i$ at the output</td>
</tr>
<tr>
<td>Span $s$</td>
<td>$p_{l,s,i}$</td>
<td>signal power of channel $i$ at the output</td>
</tr>
<tr>
<td>on Link $l$</td>
<td>$n_{l,s,i}$</td>
<td>noise power of channel $i$ at the output</td>
</tr>
</tbody>
</table>
### Appendix B. Notations

#### Notations for parameters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{L}$</td>
<td>Set of links</td>
</tr>
<tr>
<td>$\mathcal{M}$</td>
<td>Set of channels</td>
</tr>
<tr>
<td>$\mathcal{R}_i$</td>
<td>Route of channel $i$</td>
</tr>
<tr>
<td>$\mathcal{M}_l$</td>
<td>Set of channels transmitted on link $l$</td>
</tr>
<tr>
<td>$P^0_l$</td>
<td>Total power target on link $l$</td>
</tr>
<tr>
<td>$L_{l,s}$</td>
<td>Fiber loss coefficient at span $s$ on link $l$</td>
</tr>
<tr>
<td>$G_{l,s,i}$</td>
<td>Gain coefficient of channel $i$ at span $s$ on link $l$</td>
</tr>
<tr>
<td>$ASE_{l,s,i}$</td>
<td>ASE noise of channel $i$ at span $s$ on link $l$</td>
</tr>
<tr>
<td>$h_{l,s,i}$</td>
<td>Span transmission for channel $i$ at span $s$ on link $l$</td>
</tr>
<tr>
<td>$T_{l,i}$</td>
<td>Link transmission for channel $i$ on link $l$</td>
</tr>
<tr>
<td>$G_{l,i}$</td>
<td>Gain value for channel $i$ on the spectral shape of the optical amplifier on link $l$</td>
</tr>
<tr>
<td>$f_{l,s}$</td>
<td>Filter loss of the optical amplifier at span $s$ on link $l$</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>Network system matrix</td>
</tr>
<tr>
<td>$\Gamma_l$</td>
<td>Link system matrix</td>
</tr>
<tr>
<td>$\hat{\Gamma}_k$</td>
<td>Stage system matrix</td>
</tr>
</tbody>
</table>
Bibliography


[38] F. P. Kelly, A. K. Maulloo, and D. Tan. Rate control for communication networks:
Shadow prices, proportional fairness and stability. *Journal of the Operational Re-


[40] G. M. Korpelevich. The extragradient method for finding saddle points and other
756, 1976.

the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*,

[42] S. Kunniyur and R. Srikant. End-to-end congestion control schemes: utility func-
tions, random losses and ECN marks. *IEEE/ACM Transactions on Networking*,


[46] A. Mecozzi. On the optimization of the gain distribution of transmission lines with
unequal amplifier spacing. *IEEE Photonics Technology Letters*, 10(7):1033–1035,
1998.

systems. In *Proceedings of the 4th WINLAB Workshop on Third Generation Wireless


[78] N. Stefanovic and L. Pavel. Link level power control of optical networks with time-
delay. In Proceedings of the 46th IEEE Conference on Decision and Control, pages


1996.

[81] C. Tan, D. Palomar, and M. Chiang. Solving nonconvex power control problems in
wireless networks: low SIR regime and distributed algorithms. In Proceedings of the
48th IEEE Global Telecommunications Conference (GLOBECOM), November 2005.


[83] O. Tonguz and F. Flood. EDFA-based DWDM lightwave transmission systems with
end-to-end power and SNR equalization. IEEE Transactions on Communications,

ceedings of the 26th IEEE Conference on Computer Communications (INFOCOM),

[85] H. Yaiche, R. Mazumdar, and C. Rosenberg. A game theoretic framework for band-
width allocation and pricing in broadband networks. IEEE/ACM Transactions on

[86] A. Yariv. Signal-to-noise considerations in fiber links with periodic or distributed


