DYNAMIC RESOURCE MANAGEMENT FOR SYSTEMS WITH CONTROLLABLE SERVICE CAPACITY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ravi Kumar

August 2015

DYNAMIC RESOURCE MANAGEMENT FOR SYSTEMS WITH CONTROLLABLE SERVICE CAPACITY

Ravi Kumar, Ph.D.
Cornell University 2015

The rise in the Internet traffic volumes has led to a growing interest in reducing energy costs of IT infrastructure. Resource management policies for such systems, known as power aware policies, are becoming increasingly important. We propose a dynamic capacity management framework for such systems to save energy costs without sacrificing quality of service. The system incurs two types of costs; a holding cost which reflects the *Quality of Service (QoS)* experienced by the users, and a cost of effort (*Utilization/Energy cost*) based on the amount of resources being used. The key challenge for the service provider is balancing these two objectives, as using more resources improves the system performance but also leads to higher utilization costs. This tension between delivering good performance and reducing energy costs is the central feature of the models considered in this work. In order to make good capacity allocation decisions for such scenarios, we formulate two queueing control problems using the Markov Decision Process framework.

We first consider the problem of service rate control of a single-server queueing system with a finite-state Markov-modulated Poisson arrival process. We show that the optimal service rate is non-decreasing in the number of jobs in the system; higher congestion levels warrant higher service rates. On the contrary, however, we show that the optimal service rate is not necessarily monotone in the current arrival rate. If the modulating process satisfies a stochastic mono-

tonicity property, the monotonicity is recovered. We examine several heuristics and show where heuristics are reasonable substitutes for the optimal control.

In the next model we consider the problem of dynamic resource allocation in the presence of different job classes. In this case, the service provider has to make two decisions; how many resources to utilize and how to allocate these resources among various job classes. These decisions, made dynamically based on the congestion levels of job classes, minimize a weighted sum of utilization and QoS related costs. We show that a priority policy is optimal for scheduling the jobs of different classes. We further show that the optimal resource utilization levels are non-decreasing in the number of jobs of each type. We extend this model to the case when server reallocations incur a switching penalty. In this case, we show that the optimal policy is hard to characterize and computing the optimal solution using the Dynamic Programming framework becomes intractable. Instead, we develop several heuristics and perform numerical studies to compare their performance.

BIOGRAPHICAL SKETCH

Ravi Kumar was born on August 7, 1981 and grew up in the northern part of India. He earned his Bachelors degree in Mechanical Engineering from the Indian Institute of Technology Delhi, India in 2003. After working for four years in the energy sector in India, he came to the United States to pursue higher studies. He obtained a Masters degree in Mechanical Engineering from the State University of New York at Buffalo in 2009. At Buffalo, he did research on estimating the motion of tumors, a concern in radiation therapy. He joined the doctoral program in Operations Research at Cornell in 2009.

To my family and my teachers.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Mark Lewis, for his support and guidance through the course of my study here. I am extremely grateful to him for giving me the time and attention I needed. I gratefully acknowledge his patience in discussing research ideas with me and his insightful feedback that have enabled me to complete this work. Throughout our collaboration, I not only benefited from his technical knowledge but also imbibed the skill of presenting my ideas in a cohesive and lucid manner. He has constantly encouraged me and given me the necessary time to develop as an independent researcher and to pursue my research interests.

I am extremely grateful to Professor Huseyin Topaloglu for being a mentor to me throughout my stay at Cornell and for serving on my special committee. I have had the opportunity to be involved in joint research with Professor Topaloglu, which proved to be enjoyable and fruitful. He has been a constant source of inspiration, encouragement and motivation. For that I deeply thank him. I would also like to thank Professor Shane Henderson for serving on my special committee. His ideas and suggestions on my dissertation work have been very helpful.

I feel a deep sense of gratitude to my advisor during my Master's studies, Professor Tarunraj Singh. It wouldn't have been possible for me to pursue research here without his initial support and encouragement.

I would like to thank my parents and sister for their constant love, support and encouragement. They have worked hard to make my life wonderful and enable me to pursue my dreams.

I would also like to thank my fellow PhD students. Kenneth Chong who always had the time to listen to my ideas and provide good feedback (and baked goodies); Gabriel Zayas-Caban for initiating weekly research meetings that helped me learn sample-path based proof techniques; Joyjit Roy for many table-tennis games and stimulating discussions on Indian politics; Felipe Tagle for the wonderful tennis sessions; and finally Aadhar Jain, Siddarth Chandrasekaran, Sriram Sivaramakrishnan and Madhur Shrivastava with whom it was always a pleasure to talk and enjoy many games of cricket and badminton.

TABLE OF CONTENTS

Biographical Sketch iii
Dedication iv
Acknowledgements v
Table of Contents vii
List of Tables ix
List of Figures x

1 Introduction **1**
1.1 Background 1
1.2 Motivation 2
1.3 Thesis Outline and Contributions 4

2 Dynamic Service Rate Control for a Single Server Queue with Markov Modulated Arrivals **8**
2.1 Introduction 8
2.2 Model Formulation 14
2.2.1 The Discounted Expected Cost Formulation 15
2.2.2 The Long-Run Average Cost Formulation 17
2.3 Structural Properties of Optimal Policies 19
2.3.1 Monotone In The Number of Customers 21
2.3.2 Monotone in the phase process 24
2.4 Numerical Study 30
2.4.1 Comparison with Heuristics 30
2.4.2 Approximation of a System with Non-homogeneous Poisson Arrivals 37
2.5 Conclusions 44

3 Dynamic Capacity Management for a Multi-Class Queue with Controllable Service Capacity **46**
3.1 Introduction 46
3.2 Preliminaries and Model Description 49
3.2.1 System with fixed allocation 51
3.2.2 System with limited flexibility 52
3.2.3 System with full flexibility 56
3.3 Structural Properties of Optimal Policies 59
3.3.1 System with limited flexibility 60
3.3.2 Two Stations 65
3.3.3 System with full flexibility 68
3.4 Comparison of Systems - Numerical Study 76
3.5 Heuristic Policies 78
3.5.1 Performance Measures for a Priority Queue 79
3.5.2 Simple Heuristics 80

3.5.3 Single Queue (SQ) Heuristic 81
3.5.4 Numerical Study 83
3.6 Conclusions 88

4 Dynamic Capacity Management for a Multi-Class Queue with Set-up Costs 89
4.1 Introduction 89
4.2 Preliminaries and Model Description 90
4.2.1 Numerical Solution 94
4.3 Model with Consolidated Switching 98
4.4 Structure of Optimal Policy for the Consolidated Switching Model 101
4.5 Numerical Study and Heuristics 109
4.5.1 Fluid Heuristics 112
4.6 Conclusions 114

A Appendix for Chapter 2 116

B Appendix for Chapter 3 123

Bibliography 133

LIST OF TABLES

2.1 Arrival Rate Parameters for Phase Transition Process in Examples 2.4.1 and 2.4.2 32
2.2 Average Cost Rates and Percentage Difference between Optimal and Heuristic Policies for Example 2.4.1. 33
2.3 Average Cost Rates and Percentage Difference between Optimal and Heuristic Policies for Example 2.4.2 35
2.4 Average Cost Rates and Percentage Difference between Optimal and Approximate NHPP Policy 42
2.5 Average Cost Rates and Percentage Difference between Optimal and Approximate NHPP Policy 43

3.1 Comparison of average costs for the three systems in Example 3.4.1 77
3.2 Average Cost Rates and Percentage Difference between the Optimal and Heuristic Policies for Example 3.5.1 84
3.3 Input parameters for Example 3.5.2 85
3.4 Average Cost Rates and Percentage Difference between the Optimal and Heuristic Policies for Example 3.5.2 85
3.5 Input parameters for Example 3.5.2 86
3.6 Estimated mean and 95% confidence interval for the long-run average costs in Example 3.5.3 87

LIST OF FIGURES

2.1 Transition structure of Phase process for Examples 2.3.1 and 2.3.2. 24
2.2 Structure of Optimal Policy for Example 2.3.1 25
2.3 Structure of Optimal Policy for Example 2.3.2 26
2.4 Comparison of Gain Values for Heuristic and Optimal Policies for Example 2.4.1 34
2.5 Comparison of Gain Values for Heuristic and Optimal Policies for Example 2.4.2 36
2.6 NHPP Policy vs Approximate MMPP policy for various congestion levels (T=7) 42
2.7 NHPP Policy vs Approximate MMPP policy for various congestion levels ($T = \frac{3\pi}{2}$) 43

3.1 Model for the Multi-Class System 49
3.2 Breakdown of total costs for the first 8 scenarios discussed in the Example 3.4.1 78

4.1 Optimal policy for medium set-up cost case discussed in example 4.2.1. Symbols represent if servers are switching away (-), towards (+) or staying (o) at station 1. Numbers below each symbol denotes the number of active servers at each station. 98
4.2 Optimal policy for medium set-up cost case discussed in example 4.2.2. Symbols represent if servers are switching away (-), towards (+) or staying (o) at station 1. Numbers below each symbol denotes the number of active servers at each station. 99
4.3 Description of policies π_1 and π_2 over time period $[\sigma, \nu]$ for proof of Theorem 4.4.3 105
4.4 Optimal policy for Example 4.5.1 (CS Model). Symbols represent if servers are switching away (-), towards (+) or staying (o) aet station 1. The number below each symbol denotes the number of active servers. 110
4.5 Performance of the consolidated switching policy 111
4.6 Evolution of the Deterministic Fluid Queue 112
4.7 Comparison of Heuristic Policies 114

CHAPTER 1

INTRODUCTION

1.1 Background

We consider several fundamental problems related to dynamic capacity management of service systems. A typical example of the kind of system consists of a service facility with S reusable and flexible resources (*Servers*). The facility provides service to jobs belonging to m different customer classes. The inter-arrival times and service requirements of jobs are stochastic with known but class dependent distributions. The job classes may differ based on demand patterns, resource requirements, sensitivity to delay, price paid for service etc. An arriving job goes to an available server for service or waits in a queue for a server to become available. The service provider constantly monitors the number of jobs of each type waiting in the system (*State*) and evaluates an objective measure of the user experience (*Quality of Service (QoS)*). This metric can be based on the average time spent by a job in the system (*Response Time*), average *holding costs* incurred per unit time, average job drop rates (*Packet Loss*), average number of jobs served per unit time (*Throughput*) etc. The goal of a service provider is to find dynamic resource allocation or scheduling policy that minimizes costs based on some predefined *QoS* metric.

In many situations, particularly in the context of computing and telecommunications systems, in addition to the *QoS* related costs, the system also incurs a cost of effort (*Utilization cost*) based on the amount of resources being used. Moreover, the service capacity is *controllable* i.e, the number of active resources can be changed (increased/decreased) relatively quickly. In computing and telecommunications systems, a natural metric for utilization cost is the energy consumed per unit time (*Power*). The key concern for the service provider here is to not only provide a good quality of service, but to do so in an energy efficient fashion. This poses an additional challenge for the service provider; utilizing more resources improves the system performance but also leads to higher energy costs. This tension between delivering a good performance and reducing utilization costs (energy) is the central feature of the models considered in this work. Efficient operational policies for these systems must address this fundamental trade-off by making two decisions; how many resources to use and how to allocate these resources among various job classes. These decisions, made dynamically based on the state of the system, minimize a weighted sum of utilization and QoS based costs incurred by the system over the entire planning horizon. We look at these systems from the paradigm of Queueing theory and use Stochastic Dynamic Programming as the main analytical tool for characterizing the optimal policies. To put this setting into context, we provide several real world scenarios in the next section where models of this nature may be useful.

1.2 Motivation

The main motivation for this work comes from service systems that arise in the areas of computing and telecommunications. Some examples where these models arise naturally are:

- *Data Centers:* The past decade has seen a rapid growth in Cloud computing ([39]) based service models for online content delivery. Such facilities provide businesses and scientific computing users with on-demand, scalable and cost effective access to computing. Users stand to benefit from these technologies because they no longer have to invest in an expensive in-house computing infrastructure and only pay for the capacity they actually use. To provide these services, cloud service providers maintain several large-scale computing infrastructures called Data Centers comprising of tens of thousands of servers and associated cooling infrastructure [54]. The combined energy consumption of these facilities and associated internet infrastructure represented 1.1 to 1.5% of the global energy consumption in 2010 and continues to grow rapidly [54]. Data centers typically experience low resource utilization levels (10-30%) with servers idling most of the time waiting for infrequent peak loads to occur. Moreover, an idling server consumes around 60 % of its peak power, thus a significant fraction of the energy consumed gets wasted (see Gandhi et al. [22]). This scenario provides a natural setting for the implementation of the models discussed here. Current hardware and software are amenable to the implementation of dynamic capacity management policies at multiple levels (see Beloglazov et al. [9]):

- At the level of data center, such policies can help answer how many servers to operate based on the experienced loads.
- At the level of a server, the relevant issue is how many processors (CPUs) to use and how to dynamically allocate them to multiple Virtual Machines.
- At the level of a processor, one could implement Dynamic Voltage Frequency Scaling (DVFS) to vary processor speeds based on the observed loads.

- *Wireless Data Transmission:* Another relevant area for this work is in the area of mobile computing (see Biyikoglu et al. [53]). Hand-held devices that transmit and receive data over a wireless channel have become ubiquitous. Since these devices are mostly battery powered, energy if often a limited resource. Policies that control the power consumption of these devices by adjusting the transmission rate in response to the number of packets waiting to be transmitted in the buffer may help improve overall user experience.
- *On Demand Computing:* Many web service providers purchase computing, network and storage capacity (compute instance) from a cloud computing service provider to host their services. These compute instances are resizable and can be quickly scaled up or down based on the requirements of the user (e.g, Amazon AWS [1], Google AppEngine [2] and Oracle PaaS [3]). Businesses pay based on the size of the instances that they buy. A revenue maximizing service provider may be interested in policies that dynamically scale compute instances based on the observed demand patterns to deliver good performance to its users in a cost effective manner. The framework discussed in this work is suitable for developing policies for such settings.

1.3 Thesis Outline and Contributions

Problems of the nature described in the previous section are often referred to as the "Service Rate Control" problem in the OR community (see George et al. [23] and Ata [6]) and "Dynamic Power Management" problem in the computer systems community (see Wierman et al. [55]). Whereas the vast majority of research in this area has focussed on models with stationary arrival rates, internet traffic handled by computer systems often display demand patterns with fluctuating arrival rates. In this situation, a service provider implementing dynamic capacity scaling is faced with an interesting dilemma; if they do not increase the capacity in anticipation of a future burst of arrivals, the users may experience heavy congestion, while an untimely increase in capacity leads to unnecessarily high utilization costs. In Chapter 2, we address this more realistic scenario and investigate the problem of service rate control for a single server queue with time-varying arrivals. We propose a framework based on the Markov modulated Poisson processes; a popular model amongst practitioners that is also relatively easy to analyze. Assuming that the goal is to minimize a combination of utilization cost and holding cost incurred per unit time, we study this problem under both the discounted and average cost optimality criterion. In either case, we characterize the structure of an optimal service rate as being monotone in the queue length for each arrival rate but **not** necessarily monotone in the arrival rates for each queue length. In particular, we show that the manner in which the process switches between the arrival rates plays an important role in determining the structure of the optimal policy. We further prove that monotonicity in the arrival rates is recovered when the transition matrix governing the MMPP is stochastically monotone. We investigate two natural schemes that provide fast and easily implementable approximate policies. Our numerical study, which provides a detailed comparison of the performance of these heuristic policies with the optimal policy, reveals some important insights. We see that while the heuristic policies perform well in some cases, the selection of which heuristic to apply in any particular case requires a careful consideration of problem parameters. In particular, one should take into account both the transition structure of the MMPP and the rates at which the process switches between different arrival rates before selecting an appropriate heuristic.

Service systems often observe demands from multiple class of customers. For example a computing facility may be dealing with requests that are highly sensitive to delay (low latency) like streaming video or voice data, while requests from e-mail and website users can tolerate higher delays. In this situation, a service provider is interested in providing differentiated service to multiple customer streams in the most energy-efficient fashion. In chapter 3, we consider the problem of how servers can be dynamically allocated in a system with multiple class of customers. The system serves demand streams that differ based on arrival rates, service requirements and holding costs. A central controller has access to a fixed number of servers and must decide how many servers to use and how to allocate their effort among multiple demand streams. Like in the earlier model, the goal is to minimize a combination of utilization cost and holding cost incurred per unit time. We study this problem under both the discounted and average cost optimality criterion. The joint capacity control and scheduling nature of the decisions makes the exact analysis of the problem very challenging. Most of the work in this area has either focussed on minimizing class dependent utilization costs with constraints on long-run average waiting times of each class (see Yang et al. [56]) or the problem of minimizing a weighted sum of packet drop rates with a constraint on long-run average utilization cost rate (power) (see Ata et al. [7]). In both situations, the problem decomposes by job classes using the *Lagrangian* method (see Altman [5]). In our work, the utilization cost penalizes the sum of resource usage across all demand streams. Since the utilization cost is non-linear, such decomposition techniques can not be applied. Even so, we are able to characterize the optimal policy fully for the case of two customer classes. We show that a $c - \mu$ based priority scheduling policy is optimal and the resource utilization is increasing in the length of each queue. As the number of classes become large, evaluating optimal policies using numerical methods becomes increasingly difficult. We develop several heuristic policies for this case and perform numerical studies to evaluate their performance. Our numerical studies show that these heuristics perform extremely well for large number of examples.

In Chapter 4, we address the scenario when the system incurs a cost penalty when resources switch between classes in addition to the utilization and holding costs. We continue to assume that the both switches and changes in resource capacity are instantaneous. Unlike the model without set-up cost, in this case it may no longer be optimal to work at exactly one station. The optimal policy also displays many peculiarities that are hard to characterize. Furthermore, the high dimensional nature of the state space makes this model computationally intractable even for small instances. However, when the set-up costs are low to medium, a model in which all servers set-up and work at exactly one station and switch collaboratively provides good performance. In this case we are able to partially characterize the scheduling policy. We also develop fluid model based heuristic policies for the more general model and conduct numerical studies to show that these heuristics perform well.

CHAPTER 2

DYNAMIC SERVICE RATE CONTROL FOR A SINGLE SERVER QUEUE WITH MARKOV MODULATED ARRIVALS

2.1 Introduction

In this chapter, we study a fundamental queueing control problem: managing the service rate of a server in the face of non-stationary arrival rates. We have a queue with an infinite buffer and a single server. Arrivals occur according to a Markov-modulated Poisson process (MMPP), which is to say that the rate of the Poisson process driving the arrivals into the system changes according to an exogenous Markov process. This exogenous Markov process is commonly referred to as the *phase modulating process*. The service times are assumed to be exponential.[1] We incur a holding cost for each job in the system and there is a cost for running the server at different service rates. Given that the state of the phase modulating process and state of the queue are known, we want to find a policy to adjust the service rate so as to minimize either the expected discounted cost or the long-run average cost rate over an infinite horizon.

Our model is motivated by the power aware transmission policies that are becoming increasingly important over the Internet and in mobile networks. The goal of such policies is to control the power consumption of wireless devices by adjusting the transmission rate in response to the number of packets waiting to be transmitted in the buffer. Due to changes in incoming and outgoing traffic through the device, it is almost always the case that the packet arrivals display non-stationarities, creating periods of bursts followed by near-complete

[1]In Kendall's notation, our queueing system is classified as MMPP/M/1

silence. Our use of an MMPP to model arrivals is intended to capture such periods of bursts and silence. An alternative model to capture the non-stationary nature of arrivals is the non-homogenous Poisson process (NHPP), but for the wireless applications we have in mind, MMPP appears to be a more suitable model of non-stationarity since these applications involve arrival rate changes occurring at random points in time, whereas an NHPP models arrival rates as a fixed function of time. Furthermore, when computing the optimal policy under an NHPP, one needs to keep track of time together with the queue length, resulting in an uncountable state space. This issue is not present when dealing with an MMPP.

Controlling queues when arrivals have varying rates poses interesting challenges. When controlling such queues, the policy in use not only needs to consider the current arrival rate, but it also needs to anticipate the arrival rate in the near future and adjust decisions accordingly. For example, if the current arrival rate is relatively low, but arrivals are expected to be more frequent in the near future, then the control policy may choose to proactively speed up the service rate to empty the system (as much as possible) before the higher arrival rate strikes. Similarly, if the current arrival rate is high and the current system load is high, the control policy may slow down the service rate in anticipation of lower arrival rates in the near future. The extent to which changes in arrival rates can be foreseen or anticipated depends on how the non-stationarity is modeled, but policies that explicitly address the non-stationarity in arrival rate are naturally expected to make better decisions than those that do not. Furthermore, the need to model arrival non-stationarities is becoming increasingly important as non-stationary queues find more use in telecommunications applications to study congestion problems on the Internet and mobile networks.

We provide a characterization of how the optimal policy depends on the queue length and the arrival rate. Throughout, we assume that the arrival intensities change according to a general continuous time Markov chain (CTMC) defined on the state space $\{1, 2, \ldots, L\}$ and the arrival rates of the MMPP are ordered such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$. In this context, the first interesting question is whether the optimal service rate is monotone in the queue length, for a fixed state of the phase modulating process. We answer this question, not too surprisingly, in the affirmative, indicating that the optimal service rate is higher as we have more jobs in the buffer, all else being equal. The second interesting question is whether the optimal service rate is monotone in the state of the phase modulating process, for a fixed queue length. Perhaps surprisingly, the answer to this question is not necessarily affirmative, indicating that the optimal service rate is not necessarily higher as the arrival rate becomes larger, all else being equal. This observation builds the intuition that the optimal policy should anticipate the arrival rate in near future. For example, even if the current arrival rate is higher, the optimal policy may choose not to serve the jobs faster because the arrival rate is expected to slow down soon after the higher arrival rate strikes. Thus, although it may be surprising to see that the optimal service rate is not necessarily monotone in the state of the phase modulating process, this non-monotonicity embodies the intuitive expectation that the optimal policy may start using faster service rates even before higher arrival rates strike or may start using slower service rates even before arrival rates slow down. Motivated by this observation, a natural question is when we can expect the optimal policy to actually be monotone in the state of the phase modulating process so that the behavior that we just mentioned is not prevalent. We give sufficient conditions under which the optimal service rate is indeed monotone in the state of the phase modulating process. These conditions are simple to check and they only depend on the structure of the CTMC driving the phase process. These structural results are not only important in providing insights but are also useful in deriving efficient approximation methods. When the phase process for the MMPP has a large number of states, computing an optimal policy using value or policy iteration may still be a difficult task. In these situations, the structural properties of the value function can be used to develop approximate dynamic programming methods and obtain approximate results efficiently (see for example Powell [44]).

We include a numerical study with two goals in mind. First, we examine when it is important to explicitly capture the non-stationary behavior of an arrival process via MMPP as opposed to using some natural heuristic like assuming the system has stationary arrivals. To implement the optimal policy, a decision maker needs to look at both the state of the queue and the state of the phase process of the MMPP while a heuristic control mechanism based only on the queue length or some fixed service rate may be easier to implement. Thus, a comparison between the two helps in determining the value of a more complex control mechanism. Second, since we mentioned the alternative of using an NHPP to capture the non-stationarity in the arrivals, we explore the possibility of using a "suitable" MMPP to approximate the control policy for a system with an NHPP with a periodic rate function. We find that our preliminary results are encouraging. This is a significant diversion from previous studies since the focus here is on computing an optimal control and not solely on evaluating performance measures.

Most of the research related to Markov-modulated queueing systems deals with performance characteristics for systems without control. Excellent overviews of this line of work can be found in the survey paper by Prabhu [45] or more recently in Gupta et al. [26]. A hierarchical scheme based on MMPP was proposed by Muscariello et al. [40] to model the data generated by Internet users. Heffess et al. [29] used MMPP to approximate a statistical multiplexer whose inputs consist of superposition of packetized voice sources and data. A two state MMPP model was proposed by Shah-Heydari [50] to model the so-called aggregate asynchronous transfer mode traffic. For more general scenarios, Frost [21] proposed a scheme to approximate a simple NHPP using an MMPP by suitably quantizing the rate function of the NHPP into a finite number of rates. Each rate corresponds to a state in the Markov modulating process and the parameters of the MMPP model can be estimated using empirical data.

There is also a rich body of literature on the subject of monotone optimal policies for the control of a single server queue in a setting similar to the one considered here but with stationary arrivals. See, for example, the classical work of Crabill [14], Lippman [36] and Stidham and Weber [52]. In the context of telecommunications systems, the existing literature addresses a more closely related problem of service rate control of queues when the job service requirements are influenced by an exogenous stochastic process[2]. Such models arise frequently in point-to-point wireless data transmission where the induced transmission rates are affected by the time varying properties of the transmission medium. Berry [10] considers a very general model for this problem under a discrete-time Markovian setting. In this work, packet arrivals follow a batch Markov process and the state of the transmission channel varies according to a secondary discrete-time Markov chain. The data buffer and transmitter are

[2]Similar to the present work, the policy for this type of model depends on both the queue length and the state of the exogenous process.

modeled using a single server queue with finite capacity. The goal of the transmitter is to minimize the average cost rate or power consumption over an infinite time horizon subject to a constraint on packet delay. Another case with a constraint on the probability of buffer overflow is also discussed in this work. The author proves several results related to the monotonicity of the optimal policy. Motivated by mobile networks, Ata and Zachariadis [7] address the problem of finding optimal service rates for multiple users that are being served by a central controller. Data gets transmitted through a time varying channel that is being modulated by a two-state continuous time Markov chain. Packet data for each user arrives based on a Poisson process and gets stored in a finite capacity queue before getting transmitted. The objective is to maximize some measure of overall quality of service subject to a constraint on the long-run average power consumption. The authors show that the optimal service rates for each user depend only on its own queue length and the state of the transmission channel. They also present a method to explicitly characterize the optimal policy for each user. To the best of our knowledge, none of the aforementioned work considers the case of a multi-state Markov-modulated arrival process with service rate control as is discussed in this Chapter.

We have organized this Chapter as follows. In Section 2.2 we give a detailed description of the model and the associated assumptions, and formulate the problem as a Markov decision process. In the average cost case we provide stability conditions that guarantee convergence of the discounted cost value function to the relative value function. In Section 2.3 we give structural results related to the optimal policy in each case. In Section 2.4, we present a detailed numerical study comparing the performance of the optimal policy with heuristic policies and present an exploratory study related to the computation of a heuristic policy for non-homogeneous Poisson arrivals using the optimal policy for a MMPP/M/1 queue. We conclude the Chapter in Section 2.5.

2.2 Model Formulation

We consider a single server queue with infinite buffer capacity and job arrivals that follow an MMPP. Each arriving job has an exponentially distributed service requirement with mean 1. The phase transition process for arrivals is an *ergodic*, finite state continuous time Markov chain with generator matrix Q. Let the state space for this process be denoted by $\mathcal{S} := \{1, 2, \ldots, L\}$. When the phase transition process is in phase $s \in \mathcal{S}$, jobs arrive to the queue according to a Poisson process with rate λ_s. Without loss of generality we assume that the states are ordered such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$. Let the number of jobs in the system (buffer state) be denoted by $n \in \mathbb{Z}^+$, where $\mathbb{Z}^+$ is the set of non-negative integers. The service rate can be changed at the times of arrivals, departures or phase transitions. Together, the union of these event times and (in a moment) the added dummy transitions due to uniformization comprise the set of decision epochs. Based on the queue length, n, and the state of the arrival process, s, the controller selects a service rate $\mu_{n,s}$ from the compact set $\mathbb{A} = [0, \bar{u}]$, $\bar{u} < \infty$. When a service rate $\mu \in \mathbb{A}$ is chosen, the system incurs a cost at the rate of $c(\mu)$ per unit time. The cost rate function, $c(\cdot)$, is defined on $\mathbb{A}$ and is assumed to be strictly convex, continuously differentiable, strictly increasing and (without loss of generality) such that $c(0) = 0$. Furthermore, a holding or congestion cost is incurred at rate $h(n)$ per unit time when the buffer state is n. The holding cost function $h(n)$ is assumed to be convex, non-decreasing in n and such that $h(0) = 0$ and $\lim_{n\to\infty} h(n) = \infty$. In the average-cost case we assume $h(\cdot)$ to be a non-decreasing and convex with polynomial rate of growth ($h(n) \leq Cn^p$ for some $C \geq 0, p \in \mathbb{Z}^+$) and again such that $h(0) = 0$ and $\lim_{n\to\infty} h(n) = \infty$. The assumption about polynomial growth rate of the holding cost function for the average cost case is required for proving the existence of a policy that incurs costs at a finite rate.

Let Π be the set of non-anticipating policies. A stationary control policy, $\pi \in \Pi$, is defined as $\pi = \{\mu(n, s) \mid n \in \mathbb{N}, s \in \mathcal{S}\}$, where $\mu(n, s)$ is the service rate to be selected when the state of the system is (n, s). The controller remains idle when the queue is empty i.e, for any policy, $\mu(0, s) \equiv 0$. Thus, given a policy π, the overall process, $X(t)$, evolves as a two dimensional continuous time Markov chain on the state space $\mathbb{X} = \{(n, s) \mid n \in \mathbb{Z}^+, s \in \mathcal{S}\}$. Our objective is to find a control policy that minimizes the discounted expected cost or average expected cost per unit time over an infinite time horizon.

2.2.1 The Discounted Expected Cost Formulation

For $x = (n, s)$ and service rate $\mu \in \mathbb{A}$, let

$$f(x, \mu) := c(\mu) + h(n).$$

Let $\{(X^\pi(t), D^\pi(t)), t \geq 0\}$ be the stochastic process representing the evolution of states and decisions under an admissible policy π. Given the initial state x and discount factor $\alpha > 0$, the $\alpha-$discounted expected cost until time t under policy π is given by

$$v^\pi_{t,\alpha}(x) := \mathbb{E}^\pi_x\left[\int_0^t e^{-\alpha u} f(X(u), D(u))du\right], \tag{2.1}$$

where $\mathbb{E}_x[\cdot] := \mathbb{E}[\cdot|X(0) = x]$. The **total discounted expected cost** of a policy π given that the initial state of the system is x, is

$$v_\alpha^\pi(x) := \lim_{t\to\infty} v_{t,\alpha}^\pi(x). \tag{2.2}$$

Notice that the costs are non-negative so $v_{t,\alpha}^\pi(x)$ is non-decreasing. Therefore, the limit defined in (2.2) exists (it might be infinite). The **optimal total discounted expected cost** is

$$v_\alpha^*(x) := \inf_{\pi\in\Pi} v_\alpha^\pi(x).$$

A policy, π^*, is **total discounted expected cost optimal** if $v_\alpha^{\pi^*}(x) = v_\alpha^*(x)$ for all $x \in \mathbb{X}$.

We apply *uniformization* in the spirit of Lippman [36] and consider the discrete time equivalent of the continuous time Markov chain described above. The uniformization rate is chosen to be $\nu := \lambda_L + \bar{\eta} + \bar{u}$, where $\bar{\eta} \geq \max\{-Q_{ss} \mid s \in \mathcal{S}\}$ is any finite rate larger than the maximum of the holding time parameters for the phase transition process.

Let $v_{\alpha,k}(n, s)$ be the minimum total α-discounted expected cost that can be obtained during the last k transitions when starting from state (n, s). Using standard arguments of Markov decision theory [11], the discrete-time finite horizon optimality equations (FHOE) for the system can be written (for each $s \in \mathcal{S}$):

$$\begin{aligned} v_{\alpha,k+1}(0, s) = \frac{1}{\alpha + \nu}\Big[&h(0) + \lambda_s v_{\alpha,k}(1, s) \\ &+ \sum_{s'=1}^{L} Q_{ss'} v_{\alpha,k}(0, s') + (\nu - \lambda_s) v_{\alpha,k}(0, s)\Big] \end{aligned} \tag{2.3a}$$

$$\begin{aligned} v_{\alpha,k+1}(n, s) = \frac{1}{\alpha + \nu} \min_{\mu\in\mathbb{A}} \Big\{&c(\mu) + h(n) + \mu v_{\alpha,k}(n-1, s) + \lambda_s v_{\alpha,k}(n+1, s) \\ &+ \sum_{s'=1}^{L} Q_{ss'} v_{\alpha,k}(n, s') + (\nu - \lambda_s - \mu)\, v_{\alpha,k}(n, s)\Big\} \quad \text{for } n \geq 1, \end{aligned} \tag{2.3b}$$

where $v_{\alpha,0}$ is assumed to be zero for each state. Note that the cost function has compact level sets. That is, $\{((n,s),\mu)|f((n,s),\mu) \leq \beta\}$ is compact for all $\beta \in \mathbb{R}$. Since the state space is discrete, we may apply Proposition 3.1 of [20] to get $v_{\alpha,k} \uparrow v_\alpha$. Moreover, v_α satisfies (2.3) with v_α replacing $v_{\alpha,k}$ on the right hand side and $v_{\alpha,k+1}$ on the left hand side. The resulting set of equations are called the *discounted cost optimality equations* (DCOE) and are stated next for later reference (for each $s \in \mathcal{S}$).

$$v_\alpha(0,s) = \frac{1}{\alpha+\nu}\left[h(0) + \lambda_s v_\alpha(1,s) + \sum_{s'=1}^{L} Q_{ss'} v_\alpha(0,s') + (\nu - \lambda_s) v_\alpha(0,s)\right] \tag{2.4a}$$

$$\begin{aligned} v_\alpha(n,s) = \frac{1}{\alpha+\nu} \min_{\mu\in\mathbb{A}} \Big\{ & c(\mu) + h(n) + \mu v_\alpha(n-1,s) + \lambda_s v_\alpha(n+1,s) \\ & + \sum_{s'=1}^{L} Q_{ss'} v_\alpha(n,s') + (\nu - \lambda_s - \mu)\, v_\alpha(n,s) \Big\} \text{ for } n \geq 1. \end{aligned} \tag{2.4b}$$

2.2.2 The Long-Run Average Cost Formulation

In this section we provide conditions under which an average cost optimal policy exists and may be computed as a limit of discounted cost optimal policies. The **long-run average cost** or **gain** of a policy π given that the initial state of the system is x, is

$$g^\pi(x) := \limsup_{t\to\infty} v_{t,0}^\pi(x)/t,$$

where $v_{t,0}$ is as defined in (2.1). The **optimal expected average cost** $g^*(x)$ is

$$g^*(x) := \inf_{\pi\in\Pi} g^\pi(x),$$

and π^* is an **average cost optimal policy** if $g^{\pi}(x) = g^*(x)$ for all $x \in \mathbb{X}$. After uniformization the average cost optimality inequalities (ACOI) (cf. [48]) are,

$$w(n,s) \geq \frac{1}{\nu} \min_{\mu \in \mathbb{A}} \Big[-g + c(\mu) + h(n) + \lambda_s w(n+1,s) + \mu w((n-1)^+, s) \\ + \sum_{s'=1}^{L} Q_{ss'} w(n,s') + (\nu - \lambda_s - \mu) w(n,s) \Big] \text{ for } n \geq 0, s \in \mathcal{S}. \tag{2.5}$$

When the solution, (w, g) to the ACOI exists, w is called a *relative value function* and $g^*(x) = g$ is the optimal long-run expected average cost for any initial state x.

A solution to the ACOI (2.5) exists under a necessary and sufficient stability condition which is provided in (2.6) below. This condition requires that the maximum available service rate is higher than the long-run average arrival rate and coincides with the one derived by Yechiali [57] for the stability of queue with Markov-modulated arrivals. However, since Yechiali used the balance equations to show the existence of a steady state distribution there is no guarantee of finite long-run average cost. This is required for the MDP formulation provided. Since the phase transition process is assumed to be *ergodic*, it has unique stationary probabilities denoted by, $\{p_1, p_2, \ldots, p_L\}$.

Proposition 2.2.1. *There exists a stationary policy ,π, under which the system is stable (steady state distribution exists) if and only if the maximum available service rate satisfies the following condition*

$$\bar{u} > \sum_{s=1}^{L} p_s \lambda_s. \tag{2.6}$$

Furthermore, the long-run average cost under this policy, $g^{\pi}(x)$, is finite and independent of the initial state x.

Proof. See Appendix.

In the next proposition, we present results related to the existence of an optimal average-cost policy.

Proposition 2.2.2. *The following hold*

1. *For $\alpha > 0$, $v_\alpha(x)$ satisfies the DCOE* (2.4)*. Moreover, any stationary policy π_α that minimizes the right side of the DCOE* (2.4) *is α-discounted expected cost optimal.*
2. *If the stability condition (2.6) holds, we have,*
   (a) *There exists a stationary long-run average expected cost optimal policy $\pi^* = \{\mu^*(n, s) \mid n \geq 1, s \in \{1, 2, \ldots, L\}\}$ that is a limit of a sequence of discounted expected cost optimal policies $\{\pi_{\alpha_k}, k \geq 1\}$. That is, $\mu^*(n, s) = \lim_{k\to\infty} \mu_{\alpha_k}(n, s)$, where $\alpha_k \downarrow 0$.*
   (b) *The long-run average expected cost of policy π^* is $g^* = \lim_{\alpha\downarrow 0} \alpha v_\alpha(x)$ for every $x \in \mathbb{X}$. Moreover, there exists a subsequence $\alpha_k \downarrow 0$ such that $\lim_{k\to\infty} w_{\alpha_k}(x) := v_{\alpha_k}(x) - v_{\alpha_k}(\mathbf{0}) = w(x)$ for a distinguished state $\mathbf{0}$ such that (w, g^*) satisfy the ACOI* (2.5).

Proof. See Appendix.

2.3 Structural Properties of Optimal Policies

In this section we derive structural results for optimal policies for both the discounted cost and the average cost criterion. In a manner similar to [23], we use following definitions to simplify the optimality equations,

$$
\begin{aligned}
y_\alpha(0,s) &= 0 \text{ for } s = 1,2,\ldots,L, \qquad (2.7)\\
y_\alpha(n,s) &= v_\alpha(n,s) - v_\alpha(n-1,s) \text{ for } n \in \mathbb{N}, s \in \mathcal{S},\\
\phi(y) &= \max_{\mu\in\mathbb{A}}\{\mu y - c(\mu)\}, \text{ and}\\
\psi(y) &= \arg\max_{\mu\in\mathbb{A}}\{\mu y - c(\mu)\},
\end{aligned}
$$

where the argmax is a singleton by the assumptions on c (cf. Section 4.3 of [6]). The definitions above yield the following simplified form of the DCOE (2.4):

$$
\begin{aligned}
v_\alpha(n,s) = \frac{1}{\alpha+\nu}\Big[&h(n) - \phi(y_\alpha(n,s)) + \lambda_s v_\alpha(n+1,s) + \sum_{s'=1}^{L} Q_{ss'} v_\alpha(n,s')\\
&+ (\nu - \lambda_s) v_\alpha(n,s)\Big] \text{ for } n \in \mathbb{Z}^+, s \in \mathcal{S}. \qquad (2.8)
\end{aligned}
$$

In order to derive structural results for an optimal discounted expected cost policy, we make use of several important properties of functions $\phi(y) = \max_{x\in\mathbb{A}}\{yx - c(x)\}$ and its associated maximizers $\psi(y) = \arg\max_{x\in\mathbb{A}}\{yx - c(x)\}$ that were introduced in the DCOE (2.4). Recall that the *conjugate* of $c(\cdot)$, $\phi(\cdot)$, is convex (cf. [12]). Moreover, $\psi(y)$ is continuous, non-decreasing and equals $\phi'(y)$ wherever the derivative exists. As described in [6], since $(c')^{-1}(\cdot)$ is well-defined, continuous and strictly increasing we have the following characterization of the function $\psi(\cdot)$

$$
\psi(y) = \begin{cases} 0 & \text{if } y \le c'(0),\\ (c')^{-1}(y) & \text{if } c'(0) < y < c'(\bar{u}),\\ \bar{u} & \text{if } y > c'(\bar{u}). \end{cases} \qquad (2.9)
$$

It may also be established (see [23]) that $\phi(\cdot)$ is continuous and non-decreasing with the following characterization

$$\phi(y) = \begin{cases} 0 & \text{if } y < 0, \\ \int_0^y \psi(x)dx & \text{if } y \geq 0. \end{cases} \tag{2.10}$$

2.3.1 Monotone In The Number of Customers

We show the intuitive result that there exists an optimal policy that is monotone in n. We note that the structural part of the result could also be proven via the event-based dynamic programming framework of Koole [33]. We provide what we believe is an equally simple proof here for completeness.

Proposition 2.3.1. *The following hold*

1. *For each $s \in \mathcal{S}$, the optimal discounted expected cost value function, $v_\alpha(n, s)$, satisfies the DCOE* (2.4) *and is a non-decreasing, convex function of n.*
2. *There exists a discounted expected cost optimal policy $\{\mu_\alpha(n, s), n \geq 1, s \in \mathcal{S}\}$ that is non-decreasing in n for each $s \in \mathcal{S}$.*
3. *Under the assumptions that the holding cost is non-decreasing and convex with polynomial rate of growth and* (2.6) *hold, there exists a long-run average optimal policy, $\{\mu(n, s), n \geq 1, s \in \mathcal{S}\}$ that is non-decreasing in n for each $s \in \mathcal{S}$.*

Proof. We use induction and the FHOE (2.3) to prove the first result. The result holds trivially for $k = 0$. For the inductive step, suppose $v_{\alpha,k}(\cdot, s)$ is non-decreasing and convex on $\mathbb{Z}^+$ for each $s \in \mathcal{S}$. Let $u_n = \mu_{\alpha,k}(n, s)$ be the optimal service rate for the $(k+1)$-stage problem when the state is (n, s). Suppose we use the potentially sub-optimal decision u_n when the state is $(n-1, s)$. The FHOE

(2.3) yield,

$$
\begin{aligned}
v_{\alpha,k+1}(n-1,s) &\leq \frac{1}{\alpha+\nu}\Big[c(u_n) + h(n-1) + u_n v_{\alpha,k}((n-2)^+, s) + \lambda_s v_{\alpha,k}(n,s) \\
&\quad + \sum_{s'=1}^{L} Q_{ss'} v_{\alpha,k}(n-1,s') + (\nu - \lambda_s - u_n)\, v_{\alpha,k}(n-1,s)\Big], \\
&\leq \frac{1}{\alpha+\nu}\Big[c(u_n) + h(n) + u_n v_{\alpha,k}(n-1,s) + \lambda_s v_{\alpha,k}(n+1,s) \\
&\quad + \sum_{s'\neq s} Q_{ss'} v_{\alpha,k}(n,s') + (\nu + Q_{ss} - \lambda_s - u_n)\, v_{\alpha,k}(n,s)\Big] \\
&= v_{\alpha,k+1}(n,s),
\end{aligned}
$$

where the second inequality follows from the induction hypothesis. Thus, $v_{\alpha,k}$ is non-decreasing for all k.

To show convexity note that by the inductive hypothesis, $y_{\alpha,k}(n+1,s) = v_{\alpha,k}(n+1,s) - v_{\alpha,k}(n,s)$ is a non-decreasing function of n for each $s \in \mathcal{S}$. Let $u_{n+1} = \mu_{\alpha,k}(n+1,s)$ be the optimal rate for the $(k+1)$-stage problem when the state is $(n+1,s)$ and $u_{n-1} = \mu_{\alpha,k}(n-1,s)$ be the optimal rate when the state is $(n-1,s)$. The DCOE (2.4) imply (for $n \geq 1$)

$$
\begin{aligned}
(\alpha+\nu) y_{\alpha,k+1}(n+1,s) &\geq h(n+1) - h(n) - u_{n+1}(y_{\alpha,k}(n+1,s) - y_{\alpha,k}(n,s)) \\
&\quad + \lambda_s y_{\alpha,k}(n+2,s) + \sum_{s'=1}^{L} Q_{s,s'} y_{\alpha,k}(n+1,s').
\end{aligned}
$$

Similarly for $y_{\alpha,k}(n,s) = v_{\alpha,k}(n,s) - v_{\alpha,k}(n-1,s)$,

$$
\begin{aligned}
(\alpha+\nu) y_{\alpha,k+1}(n,s) &\leq h(n) - h(n-1) + (\nu - \lambda_s) y_{\alpha,k}(n,s) \\
&\quad - u_{n-1}(y_{\alpha,k}(n,s) - y_{\alpha,k}(n-1,s)) + \lambda_s y_{\alpha,k}(n+1,s) \\
&\quad + \sum_{s'=1}^{L} Q_{s,s'} y_{\alpha,k}(n,s').
\end{aligned}
$$

Using the definitions of $\nu = \lambda_L + \bar{\eta} + \bar{u}$ and $\bar{\mathbf{Q}} = \bar{\eta}\mathbf{I} + \mathbf{Q}$ we have,

$$\begin{aligned}(\alpha + \nu)(y_{\alpha,k+1}(n+1, s) - y_{\alpha,k+1}(n, s)) &\geq h(n+1) - 2h(n) + h(n-1) \\ &\quad + (\lambda_L + \bar{u} - \lambda_s - u_{n+1})\left(y_{\alpha,k}(n+1, s) - y_{\alpha,k}(n, s)\right) \\ &\quad + u_{n-1}(y_{\alpha,k}(n, s) - y_{\alpha,k}(n-1, s)) \\ &\quad + \lambda_s\left(y_{\alpha,k}(n+2, s) - y_{\alpha,k}(n+1, s)\right) \\ &\quad + \sum_{s'=1}^{L} \bar{Q}_{s,s'}\left(y_{\alpha,k}(n+1, s') - y_{\alpha,k}(n, s')\right) \\ &\geq 0.\end{aligned}$$

The second inequality follows as $h(n)$ is convex, the coefficients of $y_{\alpha,k}$ terms are non-negative and the inductive hypothesis. So $y_{\alpha,k}(\cdot, s)$ is non-decreasing on $\mathbb{Z}^+$ for all $s \in \mathcal{S}$ as required. Taking limits as $k \to \infty$ yields that v_α is non-decreasing and convex; the first result is proven.

Since the function $\psi(\cdot)$ is non-decreasing and $\mu_\alpha(n, s) = \psi(y_\alpha(n, s))$, we conclude that there exists an optimal policy for the discounted cost problem that is monotonically nondecreasing in the queue length for each $s \in \mathcal{S}$. This is the second result.

For the third result, consider a subsequence of discount factors $\{\alpha_i, i \geq 0\}$ such that $\alpha_i \to 0$ and corresponding discounted cost optimal policies $\mu_{\alpha_i}(\cdot, \cdot)$ that converge to an average cost optimal policy $\mu(\cdot, \cdot)$ (see Proposition 2.2.2). The previous result implies that for each fixed $s \in \mathcal{S}$, $\mu_{\alpha_i}(n, s) \leq \mu_{\alpha_i}(n+1, s)$. Thus, the same inequality holds for $\mu(\cdot, \cdot)$. ■

We remark that we have explicitly used the fact that argmax in ψ is a singleton (which follows from the strict convexity assumption on $c(\cdot)$). When the convexity is not assumed to be strict, the results still hold, but we need to take care to define ψ as the minimal element of the argmax and consider a subsequence of discount factors such that $w_{\alpha_i}(x) = v_{\alpha_i}(x) - v_{\alpha_i}(\mathbf{0}) \to w(x)$. Since $v_{\alpha_i}(n, s)$ is non-decreasing in n, so is w and proof in the average cost case follows in the same way as the discounted cost case except that we use the ACOI instead of the DCOE.

2.3.2 Monotone in the phase process

Since the states of the phase transition process are ordered such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$, one might conjecture that the optimal policy is non-decreasing in the phase state, s, for each congestion level, n. However, we present two examples to show that depending on the transition structure of the phase process, this property may *not* hold. In both examples, we use value iteration with $\alpha = 0.05$ to compute the optimal policy numerically. We consider an exponential cost rate function, $c(\mu) = e^{\mu} - 1$, and a linear holding cost function, $h(n) = n$. The set of permissible service rates is $\mathbb{A} = [0, 5]$.

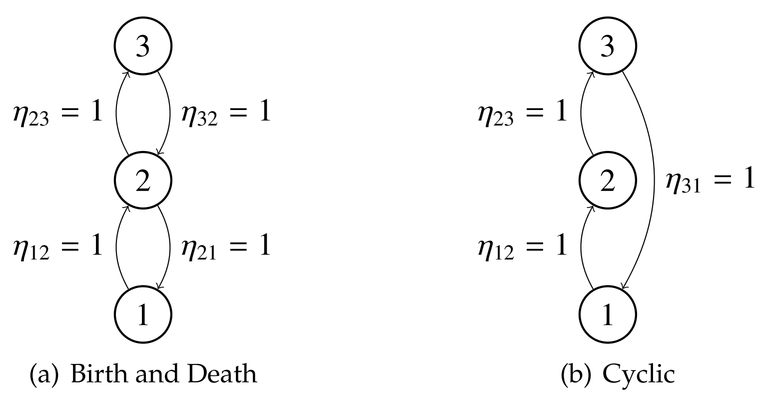


Figure 2.1: Transition structure of Phase process for Examples 2.3.1 and 2.3.2.

Example 2.3.1. In this example, the phase process is a birth and death process on the states $\{1, 2, 3\}$. See Figure 2.1(a). The infinitesimal generator for the phase

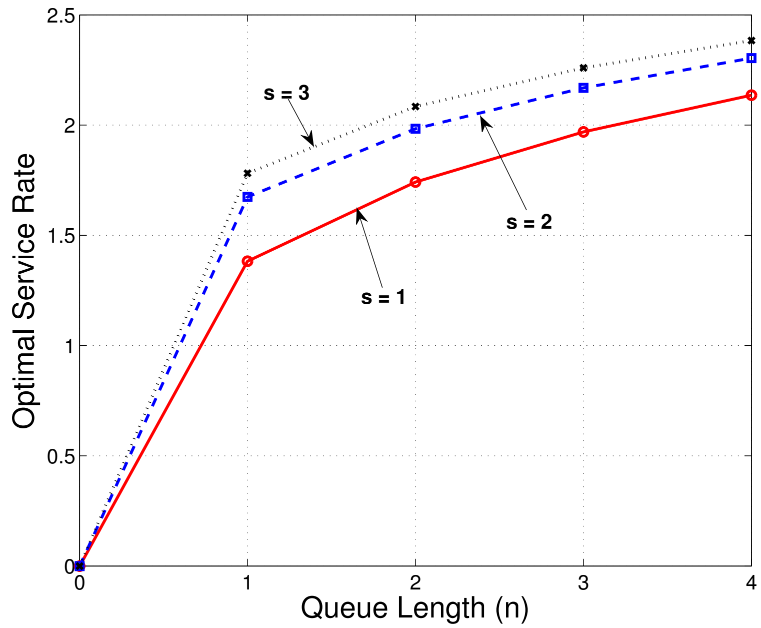


Figure 2.2: Structure of Optimal Policy for Example 2.3.1

process is given by

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix},$$

and arrival rates are $\lambda_1 = 0.5$, $\lambda_2 = 1$ and $\lambda_3 = 1.25$.

Figure 2.2 shows that the optimal policy is a non-decreasing function of n for each s. It should also be clear that the optimal service rates are non-decreasing in s for each n.

Example 2.3.2. Consider a phase process with a cyclic transition structure on the set of states $\{1, 2, 3\}$. See Figure 2.1(b). The infinitesimal generator matrix for

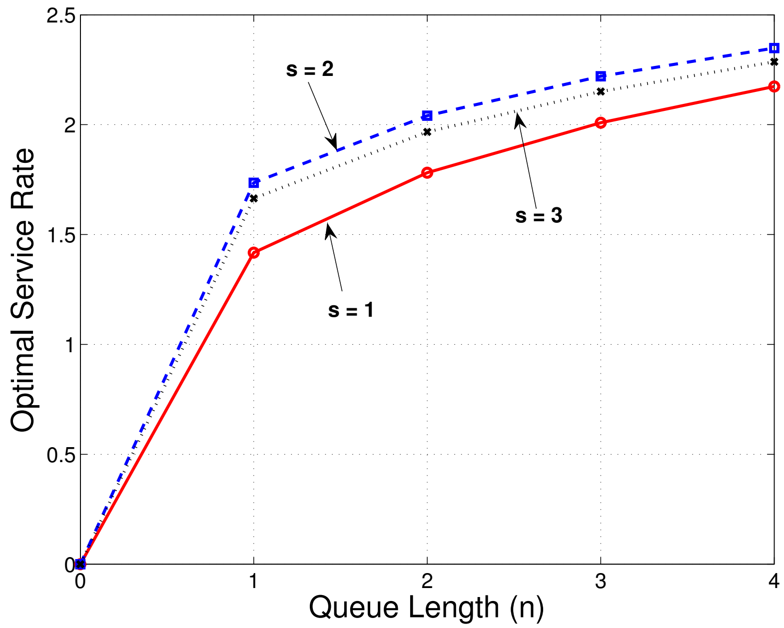


Figure 2.3: Structure of Optimal Policy for Example 2.3.2

the phase process is

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

and the arrival rates are $\lambda_1 = 0.5, \lambda_2 = 1$ and $\lambda_3 = 1.25$. Figure 2.3.1 shows that the optimal policy is non-decreasing in n for each s. However, it is clear from this figure that the service rates are not monotone in s when queue length is 4.

In Example 2.3.2, the phase process transitions from the highest arrival intensity state, 3, to the lowest arrival intensity state, 1. This causes the optimal service rate to be higher in state 2 as compared to state 3 for some congestion levels and thereby renders an optimal policy that is not monotone in s. These examples beg the question, is there a reasonable assumption under which the optimal policy is monotone in s? *Stochastic monotonicity* of the phase transition process is one such assumption.

Stochastic Monotonicity for Continuous Time Markov Chains

Intuitively, stochastic monotonicity means that given the arrival process is in a high arrival intensity state, the future states it will encounter are in some sense worse (in terms of arrival intensity) than if the process is in a low arrival intensity state. This leads to the following definitions (see for example Keilson and Kester [32]).

Definition 2.3.2. *Given two probability vectors* $\boldsymbol{p}$ *and* $\boldsymbol{q}$*, a stochastic matrix* $\boldsymbol{M}$ *and a homogeneous Markov chain* $\{X(t), t \geq 0\}$ *with probability transition function* $\boldsymbol{P}(t) = \boldsymbol{P}_{ij}(t)$*.*

1. $\boldsymbol{p}$ *stochastically dominates* $\boldsymbol{q}$ ($\boldsymbol{p} \geq_{st} \boldsymbol{q}$) *iff* $\sum_{i=n}^{N} p_i \geq \sum_{i=n}^{N} q_i$, $n = 1, 2, \ldots$.
2. *Letting* $\mathbf{M}^i$ *denote the* i^{th} *row of the matrix,* $\boldsymbol{M}$ *is called stochastically monotone if* $\mathbf{M}^k \geq_{st} \mathbf{M}^l$*, whenever* $k > l$*.*
3. $\{X(t), t \geq 0\}$ *is said to be stochastically monotone if* $\boldsymbol{P}(t)$ *is monotone.*

Note that the transition structure shown in Example 2.3.1 is stochastically monotone while that for Example 2.3.2 is not (this is trivial to see once the underlying Markov Chain is uniformized). Some useful stochastic processes that have the stochastic monotonicity property include the birth-death process, the simple random walk, the age of renewal process with decreasing failure rate [32]. In particular, the simple 2-state MMPP fluctuating between high arrival rate and low arrival rate considered by Gupta [26] and Shah-Heydari [50], is also stochastically monotone. The following provides alternative methods for specifying when a transition matrix is stochastically monotone (again refer to Keilson and Kester [32]).

Proposition 2.3.3. *For monotone matrices the following are equivalent.*

1. ***M** is monotone.*
2. $(\mathbf{T}^{-1}\mathbf{M}\mathbf{T})_{ij} \geq 0$ *where **T** is a square matrix with 1′s on or below the diagonal.*
3. $\boldsymbol{pM} \geq_{st} \boldsymbol{qM}$ *for all probability vectors* $\boldsymbol{p}, \boldsymbol{q}$ *with* $\boldsymbol{p} \geq_{st} \boldsymbol{q}$.
4. $\boldsymbol{Mv}$ *is non-decreasing for all non-decreasing vectors* $\boldsymbol{v}$.

Since for a continuous-time Markov chain with generator matrix $\mathbf{Q}$, stochastic monotonicity implies that the generator for the phase process satisfies the property $(\mathbf{T}^{-1}\mathbf{Q}\mathbf{T})_{ij} \geq 0, i \neq j$ where $\mathbf{T}$ is a square matrix with 1′s on or below the diagonal([32]). Choosing a uniformizing constant $\bar{\eta}$ yields $(\mathbf{T}^{-1}(\bar{\eta}\mathbf{I} + \mathbf{Q})\mathbf{T}) \geq 0$ in all elements. Thus using the second and fourth parts of Proposition 2.3.3 we have that $\mathbf{Q}$ is such that $\bar{\mathbf{Q}} := \bar{\eta}\mathbf{I} + \mathbf{Q}$ satisfies the property that $\bar{\mathbf{Q}}\mathbf{v}$ is non-decreasing for all non-decreasing vectors $\mathbf{v}$. This leads to the next result; the main result of this section.

Theorem 2.3.4. *Suppose that the phase transition process is stochastically monotone.*

1. *For each* $n \in \mathbb{Z}^+$, $y_{\alpha,k}(n, s)$ *is non-decreasing function of* s.
2. *There exists a discounted cost optimal policy,* $\mu_\alpha(n, s)$, *that is non-decreasing in* s *for each* n.
3. *Under the assumptions that the holding cost is non-decreasing and convex with polynomial rate of growth and* (2.6) *hold, there exists an average cost optimal policy* $\mu(n, s)$, *that is non-decreasing in* s *for each* n.

Proof. We show the first result by induction. The second and third results follow in an analogous manner to Proposition 2.3.1. The statement holds trivially for $k = 0$. Assume it holds for k. Using the definitions $\nu = \lambda_L + \bar{\eta} + \bar{u}$ and $\bar{\mathbf{Q}} = \bar{\eta}\mathbf{I} + \mathbf{Q}$ we have for $s > 1$

$$
\begin{aligned}
&(\alpha + \nu)\left(y_{\alpha,k+1}(n+1, s) - y_{\alpha,k+1}(n+1, s-1)\right) \\
&\quad = \phi\left(y_{\alpha,k}(n, s)\right) - \phi\left(y_{\alpha,k}(n, s-1)\right) + \lambda_s\left(y_{\alpha,k}(n+2, s) - y_{\alpha,k}(n+2, s-1)\right) \\
&\qquad + \left(\lambda_s - \lambda_{s-1}\right)\left(y_{\alpha,k}(n+2, s-1) - y_{\alpha,k}(n+1, s-1)\right) \\
&\qquad + \left(\lambda_L - \lambda_s\right)\left(y_{\alpha,k}(n+1, s) - y_{\alpha,k}(n+1, s-1)\right) \\
&\qquad + \left[\bar{u}\left(y_{\alpha,k}(n+1, s) - y_{\alpha,k}(n+1, s-1)\right)\right. \\
&\qquad \left. - \left(\phi(y_{\alpha,k}(n+1, s)) - \phi\left(y_{\alpha,k}(n+1, s-1)\right)\right)\right] \\
&\qquad + \Big[\sum_{s'=1}^{L} \bar{Q}_{s,s'} y_{\alpha,k}(n+1, s') - \sum_{s'=1}^{L} \bar{Q}_{s-1,s'} y_{\alpha,k}(n+1, s')\Big]
\end{aligned} \tag{2.11}
$$

Note that the inductive hypothesis implies that the first four terms in the RHS of (2.11) are non-negative. Now as $\phi(y) = \int_0^y \psi(x)dx$ and $\psi(y) \leq \bar{u}$,

$$
\begin{aligned}
\phi(y_{\alpha,k}(n+1, s)) - \phi(y_{\alpha,k}(n+1, s-1)) &= \int_{y_{\alpha,k}(n+1,s-1)}^{y_{\alpha,k}(n+1,s)} \psi(x)dx \\
&\leq \bar{u}(y_{\alpha,k}(n+1, s) - y_{\alpha,k}(n+1, s-1)).
\end{aligned}
$$

So the next to last term in RHS of (2.11) is non-negative. Furthermore, since the inductive hypothesis and the assumption on $\bar{Q}$ implies the last term in the RHS of (2.11) is non-negative. Thus, we conclude from the induction hypothesis that

$$
(\alpha + \nu)(y_{\alpha,k+1}(n+1, s) - y_{\alpha,k+1}(n+1, s-1)) \geq 0,
$$

as desired. ∎

2.4 Numerical Study

This section provides two insights using numerical examples. First, we compare the optimal control policy with two natural heuristics. When the environment is changing, it seems a decision-maker that is not armed with the current research might take one of two courses. (S)he might choose to ignore the state change of the environment altogether, or she might treat each state change as permanent and react accordingly. In either case, the resulting control policies are heuristics when compared to the optimal control that takes into account both the phase and queue length processes. The first goal is then to compare the optimal policy with these heuristics. Our results show that the fluctuation rate of the modulating process can play an important role in deciding between these heuristics. Second, as alluded to in Section 2.1, the current model can act as a heuristic itself when compared to a model with NHPP arrivals. We analyze when this is a reasonable approximation.

2.4.1 Comparison with Heuristics

In this section we present a comparison of the performance of optimal policy with several heuristics. Furthermore, we compare the optimal cost achievable with state-dependent service rates with the optimal cost achievable when the service rate is fixed for all states. As mentioned in George and Harrison [23], the difference between these costs represent the economic value of a responsive mechanism. The first heuristic that we consider uses the optimal control for an average cost problem where the arrival process is Poisson with the long-run mean arrival rate of the MMPP , $\bar{\lambda} := \sum_{s=1}^{L} p_s \lambda_s$. When applied to the original model, this policy is a function of the queue length only. We call this heuristic the *Average Rate Method* (ARM).

Since the state of the arrival process is known, the decision-maker may solve the stationary model with each potential arrival rate and change the service rate according to the current state of the arrival process. That is to say, a second heuristic is derived in the following way:

1. Compute the service rate control average cost optimal policy, π_s^h, for a system with Poisson arrivals with rate λ_s for each intensity level $s \in \mathcal{S}$.
2. The heuristic for the Markov-Modulated queue is obtained by using π_s^h when the state of the process is (n, s).

This heuristic is referred to as the *Phase Rate based Method* (PRM). Note that the long-run average arrival rate used in computing the ARM policy is influenced by both the infinitesimal generator matrix and the arrival rates of the phase process while the PRM policy relies only on the arrival rates.

When the service rate is fixed for all states (open loop policy), the queue operates as an MMPP/M/1 queue. For a given service rate, the transition matrix and corresponding steady state distribution can be computed numerically. Based on the steady state distribution one can determine the long-run average cost corresponding to that service rate. We then use a 1-D search procedure to find the service rate that minimizes long-run average cost. This is called the *Fixed Rate* policy.

A numerical study comparing the performance of these heuristics for various test cases is provided in Examples 2.4.1 and 2.4.2. In all cases, the policies and average cost are computed using value iteration where the queue length is truncated at 50. We use the cost rate function $c(\mu) = e^{\mu} - 1$ and holding cost rate $h(n) = n$. Service rates are allowed to be chosen from $\mathbb{A} = [0, 15]$. In each case the arrival rates change in accordance with the phase state $\{1, 2, \ldots, 8\}$ with the arrival rates as shown in Table 2.1.

	Arrival Rate in Phase State							
Case	1	2	3	4	5	6	7	8
I	0.1	0.35	0.6	0.85	1.1	1.35	1.6	1.85
II	0.1	0.6	1.1	1.6	2.1	2.6	3.1	3.6
III	0.1	0.85	1.6	2.35	3.1	3.85	4.6	5.35

Table 2.1: Arrival Rate Parameters for Phase Transition Process in Examples 2.4.1 and 2.4.2

Example 2.4.1. Suppose that phase process is a birth and death process on states $\{1, 2, \ldots, 8\}$ (recall Figure 2.1(a)). Fix $c > 0$. The transition rates for the phase process are $\eta_{i,i+1} = \eta_{i,i-1} = c$ for $2 \leq i \leq 7$, $\eta_{1,2} = c$ and $\eta_{8,7} = c$. A higher value for c means that the phase process transitions faster between the arrival phases. We refer to c as the *fluctuation rate scaling parameter*. For numerical analysis, we consider three sets of arrival rates for the phase process as shown in Table 2.1. For each set, the parameter c takes values $0.25, 0.50, 0.75$ and 1.00 resulting in a total of 12 different scenarios for the arrival process. Table 2.2 shows the results for all heuristics.

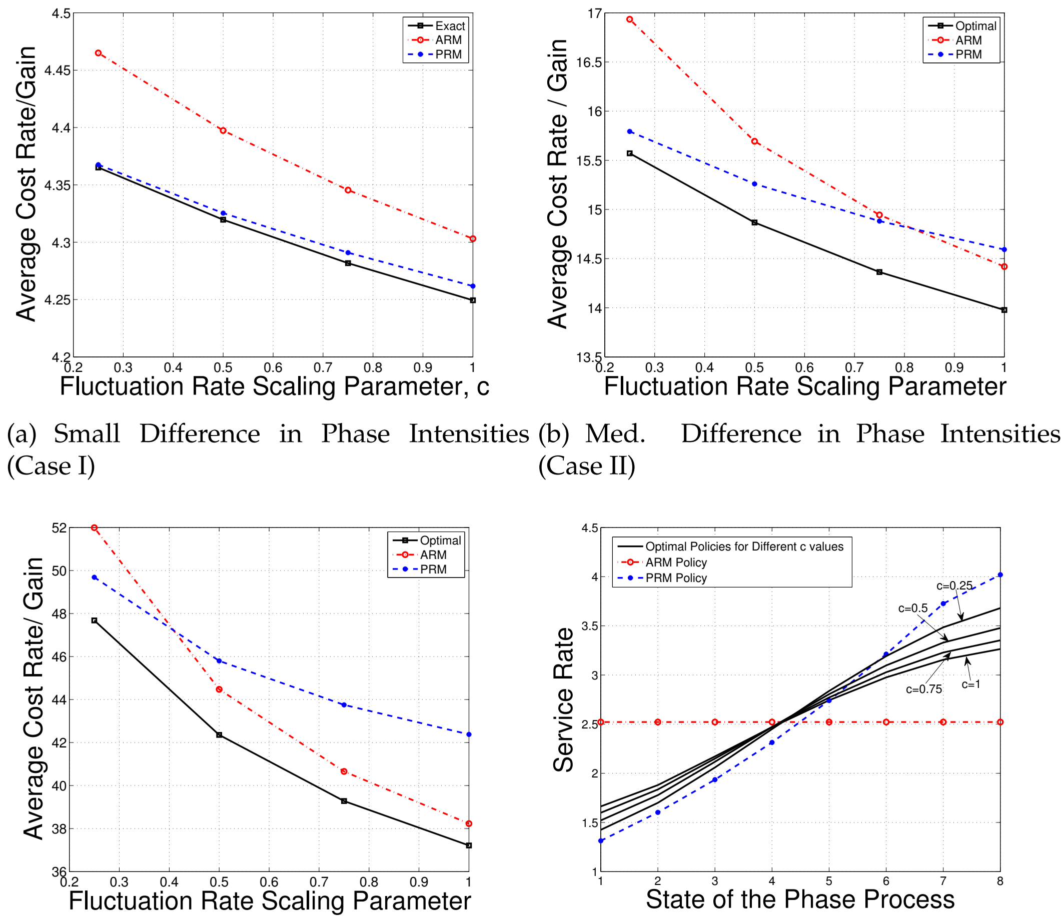
A few observations are in order. Ignoring the dynamic state information of the phase process (and using the ARM policy) is more costly when the fluctuation parameter is lower. This stands to reason since the phase process can be in a state for a long period of time, while the ARM policy assumes the arrival rate is the mean arrival rate. In Case III, when the arrival rate change is the most between phase states and with the slowest rate of changing states, the per-

Scenarios		Gain (% Sub-Optimal)			
Arrival Rates	c	Optimal	ARM	PRM	Fixed Rate
Case I	0.25	4.3651	4.4650 (2.29 %)	4.3676 (0.06 %)	7.6841 (76.03 %)
	0.50	4.3196	4.3974 (1.80 %)	4.3254 (0.13 %)	7.3185 (69.43 %)
	0.75	4.2818	4.3455 (1.49 %)	4.2909 (0.21 %)	7.0223 (64.01 %)
	1.00	4.2494	4.3031 (1.27 %)	4.2618 (0.29 %)	6.8399(60.96 %)
Case II	0.25	15.5713	16.9349 (8.76 %)	15.7936 (1.43 %)	24.8200 (59.41 %)
	0.50	14.8674	15.6939 (5.56 %)	15.2599 (2.64 %)	22.5509 (51.68%)
	0.75	14.3638	14.9444 (4.04 %)	14.8821 (3.61 %)	21.1000(46.9%)
	1.00	13.9776	14.4189 (3.16 %)	14.5924 (4.40 %)	20.1360(44.06%)
Case III	0.25	47.6797	51.9918 (9.04 %)	49.6854 (4.21 %)	61.1588(28.27%)
	0.50	42.3561	44.4741 (5.00 %)	45.7978 (8.13 %)	55.8678 (31.9%)
	0.75	39.2816	40.6579 (3.51 %)	43.7541 (11.39 %)	51.8600 (32.02%)
	1 .00	37.2150	38.2310 (2.73 %)	42.3809 (13.88 %)	48.9160 (31.44%)

Table 2.2: Average Cost Rates and Percentage Difference between Optimal and Heuristic Policies for Example 2.4.1.

cent sub-optimality for ARM is above 9%. If we try to approximate the state changes with stationary processes (using PRM) we see that again, the percent sub-optimality is high (above 13%) but this time when the fluctuation parameter is highest.

Figures 2.4(a)-2.4(c) show the change in the average cost rate under the heuristic and optimal policies as a function of the parameter c for the three arrival cases. Figure 2.4(d) shows a comparison of the heuristics and optimal policies for the arrival rates in Case III and $n = 2$ for various values of fluctuation rate parameter (the behavior is similar for other values of n). It should be clear that the PRM policy outperforms the ARM policy in most cases except for high values of c in Case III. Moreover, one should note that the performance of the ARM policy improves while that of PRM policy degrades in comparison to the optimal policy as the fluctuation rate parameter increases. Intuitively this seems reasonable since for low values of c the phase process spends more time in each phase. Therefore the PRM policy, that in each phase applies the optimal



(a) Small Difference in Phase Intensities (Case I)

(b) Med. Difference in Phase Intensities (Case II)

(c) Large Difference in Phase Intensities (Case III)

(d) Heuristic Vs Opt. Policy (Case III, $n = 2$)

Figure 2.4: Comparison of Gain Values for Heuristic and Optimal Policies for Example 2.4.1

.

policy for a stationary M/M/1 queue with arrival rate of that particular phase, performs better than the ARM policy. In fact, if $c = 0$, the PRM policy is optimal since the phase process is stationary with the arrival rate of initial phase.

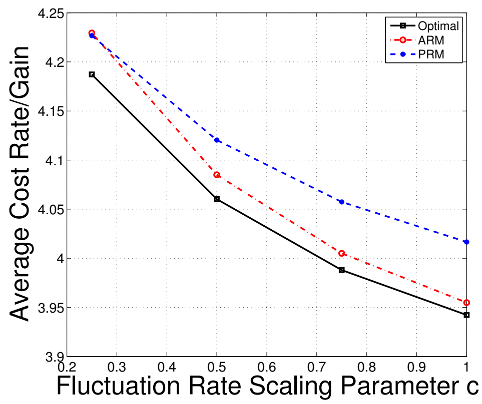
At high values of c (since the system sees more and more transitions), the arrival process behaves like a Poisson process with the average arrival rate of MMPP. Therefore the PRM policy gets penalized more in comparison to the ARM policy in this case. In Figure 2.4(d) we also see that the change in the optimal service rates as a function of phase state for each congestion level is lower for higher values of c.

Example 2.4.2. In this example we study a cyclic phase process (cf. Figure 2.1(b)) on the states $\{1, 2, \ldots, 8\}$. The transition rates for the phase process are $\eta_{i,i+1} = c$ for $1 \le i \le 7$ and $\eta_{8,1} = c$. Similar to the previous example we perform the numerical analysis for 12 scenarios for the phase process; three different sets of arrival rates given in Table 2.1 and for each set of arrival rates, c takes values $0.25, 0.50, 0.75$ and 1.00.

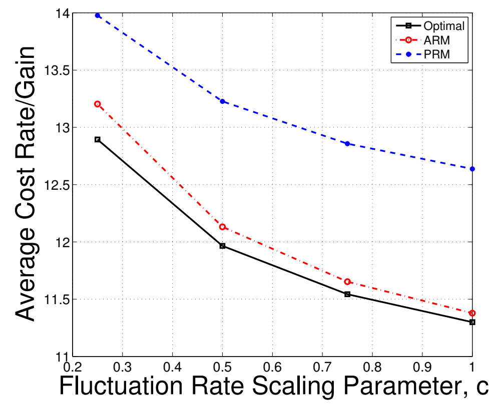
Scenarios		Gain (% Sub-Optimal)			
Arrival Rates	c	Optimal	ARM	PRM	Fixed Rate
Case I	0.25	4.1872	4.2295 (1.01 %)	4.2267 (0.94 %)	6.3440 (51.51%)
	0.50	4.0603	4.085 (0.61 %)	4.1204 (1.48 %)	5.9620 (46.84%)
	0.75	3.988	4.0051 (0.43 %)	4.0574 (1.73 %)	5.7700 (44.68%)
	1.00	3.9423	3.9549 (0.32 %)	4.0166 (1.89 %)	5.6647 (43.69%)
Case II	0.25	12.894	13.2042 (2.41 %)	13.9767 (8.39 %)	17.2439 (33.74 %)
	0.50	11.9656	12.1319 (1.39 %)	13.2268 (10.54 %)	15.6149 (30.5 %)
	0.75	11.5435	11.6531 (0.95 %)	12.8573 (11.38 %)	14.9350 (29.38 %)
	1.00	11.2996	11.3786 (0.70 %)	12.6374 (11.84 %)	14.51 (28.43 %)
Case III	0.25	31.2724	32.1887 (2.93 %)	39.4752 (26.23 %)	35.5711 (13.75 %)
	0.50	28.3046	28.7893(1.71 %)	37.1449 (31.23 %)	33.8800 (19.7 %)
	0.75	27.0506	27.3664 (1.16 %)	36.0660 (33.33 %)	32.7185 (20.95%)
	1.00	26.3445	26.5702 (0.86 %)	35.4401 (34.53 %)	31.9436 (21.25 %)

Table 2.3: Average Cost Rates and Percentage Difference between Optimal and Heuristic Policies for Example 2.4.2

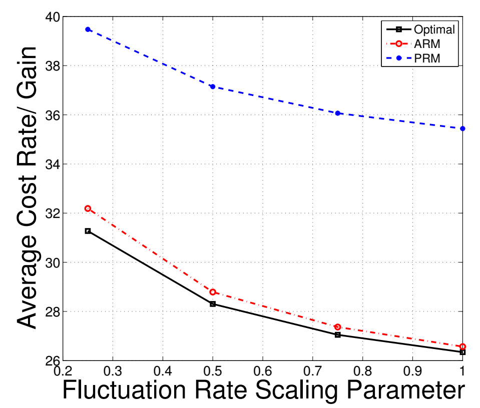
Figures 2.5(a)-2.5(c) show the change in the average cost computed under the heuristics as well as the optimal policies as a function of parameter c for the three arrival cases. It is interesting to observe that unlike Example 2.4.1, the ARM policy outperforms the PRM policy in almost all cases. As illustrated in Example 2.3.2, when the phase process has cyclic transitions, the optimal service rates may not be monotone in the phase process (for each fixed n). In fact, while



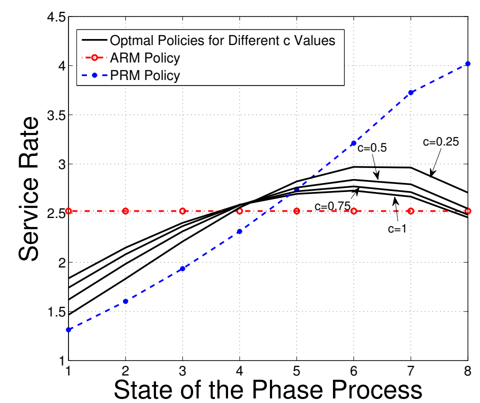
(a) Small Difference in Phase Intensities (Case I)



(b) Med. Difference in Phase Intensities (Case II)



(c) Large Difference in Phase Intensities (Case III)



(d) Heuristic Vs Optimal Policy (Case III, $n = 2$)

Figure 2.5: Comparison of Gain Values for Heuristic and Optimal Policies for Example 2.4.2

the service rates for the PRM policy are monotone in the phase of the transition process for each congestion level, the service rates in the optimal policy may begin to decrease as the phase state increases. This can be more clearly observed in Figure 2.5(d) which shows a comparison of heuristic and optimal policies as a function of the phase state when $n = 2$ for various values of parameter c and the arrival rates of Case III. Thus the ARM policy approximates the optimal policy better than the PRM policy which explains the observed performance difference.

Similar to the previous example, we observe that the performance of the ARM policy gets worse and that of PRM policy improves with a decrease in the fluctuation rate parameter c. Furthermore, the average cost percent differences provided in Table 2.3 show that the ARM policy performs extremely well for all three arrival rate cases (less than 5% from optimal for all cases). The PRM policy performs well for Case I but the degradation in its performance is quite significant (almost 35 % from optimal for Case III with $c = 1$) when the difference in arrival rates is medium or high (Cases II and III).

Tables 2.2 and 2.3 show the optimal costs achievable under the fixed rate mechanism under the column heading "fixed rate". We find that costs are between 13% and 76% suboptimal when using a fixed rate mechanism relative to the variable rate mechanism. This shows that there is substantial benefit in investing in a responsive mechanism. Furthermore, these examples show that one cannot rely on a particular heuristic method to perform well in all scenarios. In particular, we find that within the gamut of simple heuristic methods considered here, the transition structure and the transition rates of the phase process play an important role in the selection of an appropriate approximation method.

2.4.2 Approximation of a System with Non-homogeneous Poisson Arrivals

When the arrival process follows known rate changes, a non-homogeneous Poisson process is a reasonable modeling tool. In the classical work of Green and Kolesar [24] or Massey and Whitt [38] the *analysis* of queues with non-stationary arrivals is considered. From the standpoint of *control*, Yoon and Lewis [58] consider the case of admission and pricing control. One thing is certain from Yoon and Lewis's work, control of non-stationary processes can be computationally intensive. This is due the fact that to solve each instance the numerical approach requires the time to be discretized.

In this section we explore the possibility of computing an approximate average cost optimal policy for a single server queue with non-homogeneous Poisson arrivals using the optimal policies for a system with a "suitable" Markov-modulated Poisson arrival process. Apart from the arrival process, other details are the same as the setting described in Section 2.2. Let the arrival process be an NHPP with rate $\lambda(t)$. Assume that $\lambda(t)$ is a periodic function with period T. Since the optimization criterion considered in this study is over an infinite time horizon, and the rate function for NHPP is a periodic function of time, the *principle of optimality* implies that only the time elapsed in the current period and the number of jobs in the system need to be included in the state space [58].

To compute the optimal policy for an NHPP arrival process numerically, the time period is divided into n equally spaced segments of length $\Delta t = T/n$. Denote the state space for this discretized process as $\mathbb{X} = \{(n, z) \mid n \in \mathbb{Z}^{+}, s \in \{0, \Delta t, \ldots, T - \Delta t\}\}$. Under this setting, the decision epochs are the time points $0, \Delta t, 2\Delta t, \ldots, T - \Delta t$. Let ν be a uniformizing rate of the process. An event (arrival, departure or dummy transition) occurs at a decision epoch with probability $1 - e^{-\nu\Delta t}$ and with probability $e^{-\nu\Delta t}$ no event occurs. The standard theory of Markov decision processes yields the following average cost optimality inequality (ACOI):

$$
\begin{aligned}
w(n,z) \geq \min_{x\in\mathbb{A}} \Big\{ & (-g + h(n) + \mathbf{1}_{\{n>0\}} c(x))\,\Delta t + (1 - e^{-\nu\Delta t})\Big(\frac{\lambda(z)}{\nu} w(n+1, z+\Delta t) \\
& + \mathbf{1}_{\{n>0\}} \frac{x}{\nu} w(n-1, z+\Delta t) + (1 - \frac{\lambda(z)}{\nu} - \mathbf{1}_{\{n>0\}} \frac{x}{\nu}) w(n, z+\Delta t)\Big) \\
& + e^{-\nu\Delta t} w(n, z+\Delta t)\Big\} \text{ for } n \in \mathbb{Z}^+, z = 0, \Delta t, \ldots, T - 2\Delta t, \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
w(n, T-\Delta t) \geq \min_{x\in\mathbb{A}} \Big\{ & (-g + h(n) + \mathbf{1}_{\{n>0\}} c(x))\,\Delta t + (1 - e^{-\nu\Delta t})\Big(\frac{\lambda(T-\Delta t)}{\nu} w(n+1, 0) \\
& + \mathbf{1}_{\{n>0\}} \frac{x}{\nu} w(n-1, 0) + (1 - \frac{\lambda(z)}{\nu} - \mathbf{1}_{\{n>0\}} \frac{x}{\nu}) w(n, 0)\Big) \\
& + e^{-\nu\Delta t} w(n, 0)\Big\} \text{ for } n \in \mathbb{Z}^+,
\end{aligned}
$$

where $\mathbf{1}_E$ is the indicator function of the event E. When the solution, (w, g) to the ACOI exists, w is called the *relative value function* and $g^*(x) = g$ is the optimal long-run expected average cost for any initial state x.

We now present a method for constructing an approximate policy for NHPP arrivals. The main idea is to approximate the NHPP by an appropriately constructed MMPP. This is done by dividing the time period T into l subintervals and constructing an MMPP with the same number of phases as the number of subintervals i.e, l. We choose a cyclic transition structure for the phase process with arrival rate in each phase as the average rate over that subinterval. Transition rates for the phase process can be selected such that the mean sojourn time in phase s is the width of the corresponding interval. The optimal policy corresponding to this MMPP can then be applied to the original NHPP arrivals. The detailed procedure to evaluate the approximate policy is described below:

1. Partition the interval $[0, T]$ into l subintervals, $[t_{s-1}, t_s]$, $s = 1, \ldots, l$ with $t_0 = 0$ and $t_l = T$.

2. Compute average rates over each partition,

$$\lambda_s = \frac{\int_{t_{s-1}}^{t_s} \lambda(t)dt}{(t_s - t_{s-1})}.$$

3. Compute transition rates, $\eta_{i,j}$, for the phase process of MMPP,

$$\eta_{i,j} = \left\{ \begin{array}{ll} \frac{1}{(t_i - t_{i-1})}, & 1 \leq i \leq l-1, j = i+1; \\ \frac{1}{(t_l - t_{l-1})}, & i = l, j = 1; \\ 0, & \text{otherwise} \end{array} \right\}$$

.

4. Compute the optimal policy corresponding to the MMPP arrival process constructed in previous steps. Denote this policy as $\mu(n, s)$, $n \in \mathbb{Z}^+, s \in 1, 2, \ldots, l$.

5. Construct the approximate policy, $\hat{\mu}$, for the NHPP process as

$$\hat{\mu}(n, t) = \mu(n, s) \text{ for } t_{s-1} \leq t \leq t_s \, , n = 0, 1, 2, \ldots.$$

A numerical study comparing the performance of the approximation procedure stated above for various test cases is provided in Examples 2.4.3 and 2.4.4. In all cases, the policies and average cost are computed using value iteration where the queue length is truncated at 50 to keep the size of the state space manageable.

Example 2.4.3. In this example, we consider a NHPP with the following (periodic) rate function,

$$\lambda(t) = \begin{cases} 0.1 & \text{for} \quad 0 \le t < T/5 \\ 2.0 & \text{for} \quad T/5 \le t < 2T/5 \\ 4.0 & \text{for} \quad 2T/5 \le t < 3T/5 \\ 2.0 & \text{for} \quad 3T/5 \le t < 4T/5, \\ 0.1 & \text{for} \quad 4T/5 \le t < T, \end{cases}$$

where T is the time period of the rate function. One can think of this rate function as a quantized version of a triangular waveform with time period T. The discretization interval Δt, for solving the problem with NHPP arrivals is selected as 0.05 units. The cost rate function is $c(\mu) = e^{\mu} - 1$ and holding cost function is $h(n) = n$. Service rates can be selected from a set $\mathbb{A} = [0, 10]$. For computing the approximate policy, we partition the interval $[0, T]$ into 5 subintervals of equal length. Thus, the arrival rates for the corresponding MMPP are $\lambda_1 = 0.1, \lambda_2 = 2.0, \lambda_3 = 4.0, \lambda_4 = 2.0$ and $\lambda_5 = 0.1$ and the associated generator matrix is

$$\mathbf{Q} = \frac{5}{T} \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix},$$

Figure 2.6 show a comparison of optimal policy with approximate MMPP policy. Table 2.4 gives the average cost percent difference between the performance of approximate and optimal policy for various test scenarios. This data shows that the approximate policies perform extremely well in all cases (less than 1% sub-optimal).

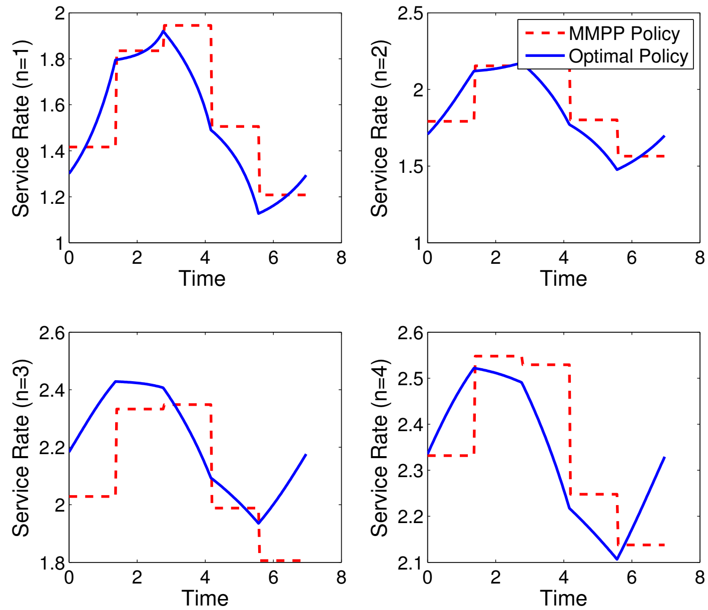


Figure 2.6: NHPP Policy vs Approximate MMPP policy for various congestion levels (T=7)

Time Period (T)	Gain(% Sub-Optimal)	
	Optimal	App. MMPP
4	8.5667	8.5932 (0.31%)
5	8.7262	8.7750 (0.41%)
6	8.7467	8.7925 (0.52%)
7	8.8225	8.8785 (0.64%)

Table 2.4: Average Cost Rates and Percentage Difference between Optimal and Approximate NHPP Policy

Example 2.4.4. In this example, we consider a NHPP with the following (periodic) rate function,

$$\lambda(t) = 5sin(\omega T) + 6$$

where T is the time period of the rate function and $\omega = \frac{2\pi}{T}$ the frequency. We consider the cases in which T is set to $\frac{n\pi}{2}, n = 1, 2, 3, \text{and}, 4$. The discretization interval Δt, for solving the problem with NHPP arrivals is selected as $\frac{T}{200}$ units.

The cost rate function is $c(\mu) = \frac{\mu^2}{2}$ and holding cost function is $h(n) = (n - 20)^+$. Service rates can be selected from a set $\mathbb{A} = [0, 15]$. For computing the approximate policy, we partition the interval $[0, T]$ into 6 subintervals of equal length.

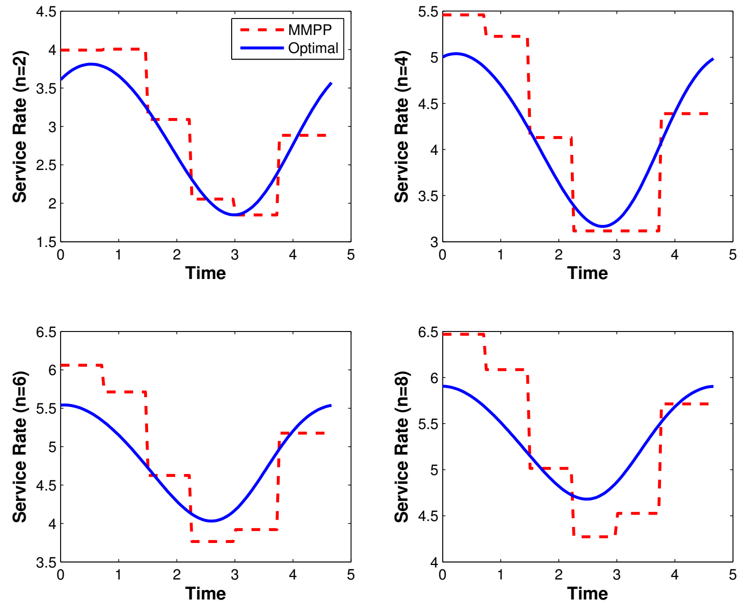


Figure 2.7: NHPP Policy vs Approximate MMPP policy for various congestion levels ($T = \frac{3\pi}{2}$)

Time Period (T)	Gain(% Sub-Optimal)	
	Optimal	App. MMPP
$\frac{\pi}{2}$	18.90	19.38 (2.52%)
π	18.91	19.32 (2.21%)
$\frac{3\pi}{2}$	18.93	19.08 (0.82%)
2π	18.94	19.29 (1.84%)

Table 2.5: Average Cost Rates and Percentage Difference between Optimal and Approximate NHPP Policy

Figure 2.7 shows a comparison of the optimal policy with heuristic MMPP policy. Table 2.5 gives the average cost percent difference between the performance of approximate and optimal policy for various test scenarios. We can again see that the approximate policies perform well in all cases (less than 3% sub-optimal).

Of course, we can not conclude from this limited analysis that MMPP approximate policies will perform well for NHPP in more general settings. Our motivation in presenting this analysis is to stimulate further research in this direction.

2.5 Conclusions

In this Chapter we investigate the problem of service rate control for a single server queue with time-varying arrivals. We propose a framework based on the Markov modulated Poisson processes; a popular model amongst practitioners that is also relatively easy to analyze. Assuming that the goal is to minimize a combination of effort cost and holding cost incurred per unit time, we study this problem under both the discounted and average cost optimality criterion. In either case, we characterize the structure of an optimal service rate as being monotone in the queue length for each arrival rate but **not** necessarily monotone in the arrival rates for each queue length. In particular, we show that the manner in which the process switches between the arrival rates plays an important role in determining the structure of the optimal policy. We further prove that monotonicity in the arrival rates is recovered when the transition matrix governing the MMPP is stochastically monotone.

There are several ways to extend our work. Our numerical study confirms that in some cases simple heuristics may perform well in the face of changing arrival rates. However, it also points out that careful selection based on the parameters of the system is required, and in many cases applying the proposed model is essential (as opposed to the heuristics). The second part of our numerical work points to the fact that our model can be used as a heuristic itself. We show that we can potentially provide a policy for a system with non-homogenous Poisson arrivals using the optimal policy for an MMPP/M/1 queue. Our results indicate that this may be a promising direction for future research.

Another problem of interest is that of the control of MMPP/M/1 queue with a finite buffer but with an explicit constraint on the job loss rate. Under this setting, while the technical conditions required for stability are not needed, handling the explicit constraint poses a significant challenge. We note that results provided by Ata [6] for the stationary arrival case may be useful.

The model under study assumes that complete information about arrival statistics is available. This may be unreasonable in situations where arrival statistics cannot be associated with the observable features of the system. A promising direction of work may be to tackle such situations using the partially observable Markov decision process framework.

Finally, we would like to point out that although for ease of exposition, we assume that the cost of effort function, $c(\mu)$, is strictly convex, continuously differentiable and non-decreasing, our proofs (with minor modification) and results hold for more general cost of effort functions. Using the analysis presented by George and Harrison [23], it can be easily shown that the structural results for an optimal policy continue to hold when $c(\mu)$ is assumed to be non-decreasing and continuous.

CHAPTER 3

DYNAMIC CAPACITY MANAGEMENT FOR A MULTI-CLASS QUEUE WITH CONTROLLABLE SERVICE CAPACITY

3.1 Introduction

In this Chapter, we consider the problem of dynamic capacity management in a service system serving multiple types of customers. Service systems often observe demands from multiple class of customers that differ based on demand patterns, resource requirements, sensitivity to delay, price paid for service etc. For example, a computing facility or a wireless data transmission node may be dealing with requests that are highly sensitive to delay (low latency) like streaming video or voice data, while requests from e-mail and website users can tolerate higher delays. The service provider constantly monitors the number of jobs of each type waiting in the system (*State*) and evaluates an objective measure of the user experience (*Quality of Service (QoS)*). Here QoS concerns are modeled using class dependent holding costs per job. The service facility has access to a limited amount of reusable and flexible resources that can be used to process the demand streams. Moreover, the service capacity is *controllable* i.e, the changes (increase/decrease) in the amount of resources being utilized and their allocation across demand streams can be made relatively quickly. Based on the amount of resources being used, the system also incurs a cost of effort (*Utilization cost*). In computing and telecommunications systems, a natural metric for utilization cost is the energy consumed per unit time (*Power*).

The key concern for the service provider here is to not only provide a good quality of service, but to do so in a cost effective manner. This poses an interesting challenge for the service provider; using more resources improves the system performance but also leads to higher utilization costs. This tension between delivering a good performance and reducing utilization costs is the central feature of the models considered in this Chapter. Efficient operational policies for these systems must address this fundamental trade-off by making two decisions; how many resources to use and how to allocate these resources among various job classes. These decisions, made dynamically based on the state of the system, minimize a weighted sum of utilization and QoS based costs incurred by the system over the entire planning horizon. We look at this problem from the paradigm of Queueing theory and use Stochastic Dynamic Programming as the main analytical tool for characterizing the optimal policies.The goal is to find a dynamic resource allocation policy to minimize either the expected discounted cost or the long-run average cost over an infinite horizon.

The joint capacity control and scheduling nature of the decisions makes the exact analysis of the problem very challenging. Most of the work in this area has either focussed on the service rate (capacity) control problem in the presence of a single class or the problem of scheduling a server (fixed capacity) in queueing systems with multiple job types. A review of the service rate control problem is provided in Chapter 2 Section 2.1. There is an extensive literature available on the problem of server scheduling in parallel queues. It is well know that the scheduling policy in multi-class queues often follows a priority policy called the $c - \mu$ rule (see for example Harrison [27],[28], Buyukkoc et al. [13], Baras et al. [8], Nain et al. [42]). According to this policy, if the holding cost for job class i is c_i and the average processing time of job is μ_i, then the policy that serves the jobs based on priorities set in decreasing order of $c_i\mu_i$ indices minimizes the holding costs. For the problem of scheduling with more than one server, Ahn et al. [4] considered a joint routing and scheduling problem with two collaborating servers, they showed that for the case when the server collaboration is additive (or super-additive) the $c - \mu$ scheduling rule is optimal. However, most of these works seek to only optimize some QoS based metric, whereas we consider a combination of QoS based cost and utilization costs.

In the area of joint capacity control and scheduling, Yang et al. [56] consider the problem of minimizing the sum of class dependent utilization costs with constraints on long-run average waiting times of each class. They characterize the structure of optimal policy under the long-run average cost criteria. Ata et al. [7] look at the problem of minimizing a weighted sum of packet job rate with a constraint on long-run average utilization cost rate (power). In both these works, the problem decomposes by job classes using the *Lagrangian* method (see Altman [5]). Lim et al. [35] consider the problem of joint service rate and admission control of two queues with class dependent routing, holding and utilization costs. They characterize the optimal policy service policy as non-decreasing in queue lengths and the routing policy as threshold type. In most of these models, the total utilization cost is the sum of class dependent utilization costs, whereas in the models considered in this dissertation, the utilization cost is incurred based on the sum of resource usage across all demand streams. Since this cost is non-linear, the dependence across stations via the utilization cost makes this problem challenging. The closest work to that of ours is of Kaufman et al. [17]. They look at the problem of joint capacity control and scheduling but for a two station tandem queueing system.

We have organized this chapter as follows. In Section 3.2 we give a detailed description of three related models for dynamic capacity management problem,

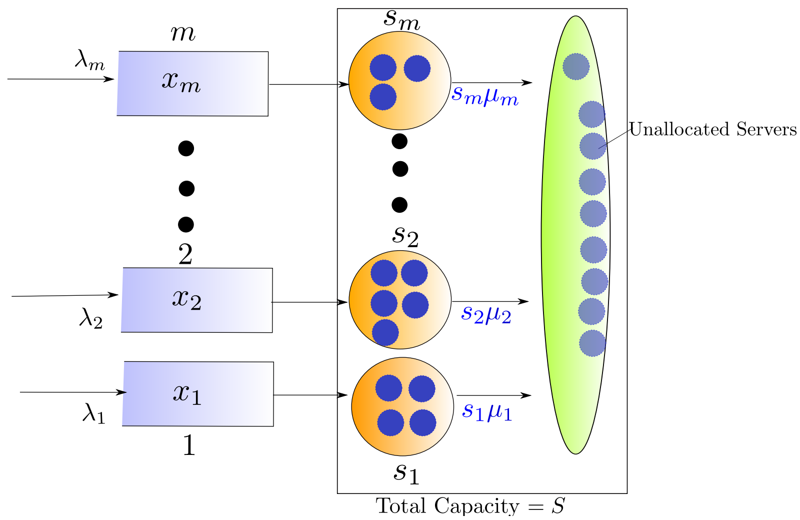


Figure 3.1: Model for the Multi-Class System

and formulate these as Markov decision processes. In Section 3.3 we give structural results related to the optimal policy for these models. In Section 3.4, we present a detailed numerical study comparing the performance of the optimal policy for the three models. We also develop several heuristic policies and perform numerical studies to evaluate their performance.

3.2 Preliminaries and Model Description

Consider a service system that processes jobs from m different job classes. Jobs of class i arrive to the system independently of other customer classes according to a Poisson process with rate λ_i, $i = 1, \ldots, m$ and wait in separate queues for service. Assume that the queues have infinite buffer size and denote the queue corresponding to job class i as station i. The service facility has access to a pool of S resources to service the jobs in the system. When s_i resources are assigned to station i, the jobs at station i are served according to the first-come-first-serve (FCFS) service discipline with rate $\mu_i(s_i)$. If one unit of capacity is assigned to station i, a job of class i requires an exponential amount of time with mean $\frac{1}{\mu_i}$ to finish service (independent of other jobs). To keep the problem tractable, we assume that the resources can collaborate perfectly i.e, $\mu_i(s_i) = s_i\mu_i$. Furthermore, we assume the preemptive resume service discipline, i.e., jobs at any station can be interrupted to make server reassignments. These are slightly restrictive assumptions since in some practical situations, the service mechanism is non-preemptive and the service rate induced by collaborating resources typically grows sub-linearly in the number of allocated resources. This is due to communication overhead incurred when resources provide service in a collaborative fashion. After receiving service at its respective station, the job leaves the system. Let $\Lambda_i = \frac{\lambda_i}{\mu_i}$, we will refer to Λ_i as the "load" at station i. Note that Λ_i is not the proportion of time that station i is busy, instead it is the amount of resources working at station i in the long-run under any stable policy. For stability we also assume that $S > \sum_{i=1}^{m} \Lambda_i$, i.e, the system has enough capacity to satisfy the requirement at all stations. When the service facility uses s resources a utilization cost of $c(s)$ per unit time is incurred. Assume that the utilization cost, $c(\cdot)$, is a continuous, convex and non-decreasing function on $A = [0, S]$ with $c(0) = 0$. Note that although we have considered that capacity can be selected from a connected set A, we will see later that this is without the loss of generality and discrete capacity sets can be modeled using piece-wise linear utilization cost functions. Furthermore, for each job waiting to be served at station i, the system incurs a holding cost of h_i per unit time. We seek resource allocation strategies that minimize the long-run average cost or discounted cost incurred over an infinite horizon.

We study three resource allocation paradigms that arise naturally for such service systems. The simplest strategy is to assign a fixed dedicated capacity to each of the stations at the start of operations. The allocated resources to each station are always on and the initial allocation is made in a way that minimizes the long-run average cost incurred by the system. In the second configuration we allow for limited flexibility of resources by assigning dedicated resources to each station at the start, but allow a central controller to dynamically choose the amount of resources to use at stations amongst the capacity assigned to them. Finally we study a system with a fully flexible configuration. Here the resources are fully shared amongst the stations and a central controller decides the amount of resource to be allocated to various stations based on currently pending jobs in the system. In both second and the third configurations, once the allocation decision is made, the capacity is immediately assigned to the stations. It should be clear that the operational costs are highest in the first configuration and lowest in the third configuration. However, from the perspective of implementation, the first system is the easiest to operate while the second and the third systems are increasingly more complex. A comparison of these systems helps gauge the value of having a more dynamic and flexible service mechanisms.

3.2.1 System with fixed allocation

In this system we fix the number of resources at each station at the start of operation. The resources allocated to each station are always on and are not shared with other stations. Let s_i be the resource level assigned to station i. Since this allocation remains unchanged, each station behaves as an independent $M/M/1$ queue. The proportion of time that station i is busy in the long-run is given by $\rho_i(s_i) = \frac{\lambda_i}{s_i\mu_i}$ and the average number of jobs at this station in long-run is $Q_i(\infty) = \frac{\rho_i(s_i)}{1-\rho_i(s_i)}$. Therefore the allocation that minimizes the long-run average cost can be computed by solving the following convex optimization problem

$$\begin{aligned}
\text{minimize} \quad & c(\sum_{i=1}^{m} s_i) + \sum_{i=1}^{m} h_i \frac{\rho_i(s_i)}{1-\rho_i(s_i)} \\
\text{subject to} \quad & \sum_{i=1}^{m} s_i \leq S \\
& s_i \geq \frac{\lambda_i}{\mu_i}, \quad i = 1, \ldots, m.
\end{aligned} \tag{3.1}$$

Note that the objective function is the sum of cost of usage and the long run holding costs incurred per unit time. The first constraint enforces the capacity limit and the remaining constraints ensure that each station is stable in the long-run.

3.2.2 System with limited flexibility

In this system station i has access to A_i dedicated resources. A central controller constantly monitors the number of pending jobs at each station and chooses to dynamically add/remove resources at station i amongst the A_i available resources at this station. We assume that the total dedicated capacity at station i is decided apriori by a system manager and satisfies $A_i > \frac{\lambda_i}{\mu_i}$ and $\sum_{i=1}^{m} A_i = S$. One way to decide this a priori capacity is to apportion the total capacity S proportionally to various stations based on their respective loads. Recall that $\Lambda_i = \frac{\lambda_i}{\mu_i}$, then under this scheme $A_i = \frac{\Lambda_i}{\sum_{k=1}^{m} \Lambda_i} S$. Note that the fixed capacity assigned to stations as per this scheme may not be optimal since we did not take into account the holding costs. A better strategy may come from the solution of the fixed allocation problem. In particular, if s^* be the solution of (3.1), we allocate

$A_i = \frac{s_i^*}{\sum_{k=1}^m s_k^*} S$ servers to station i. Note that although we have a fixed number of servers set-up at station i, the queues don't behave as independent M/M/1 queues as we still incur utilization costs based on the total amount of resources used across all stations. We formulate this problem as a Markov decision process.

Let $x_i(t)$ denote the total number of jobs at station i (including the job being processed)at time t. The state space is then given by $\mathbb{X} = \{x = (x_1, \ldots, x_m) | x \in Z_+^m\}$. At time t, the central controller chooses the amount of resources to be used at each station i, $s_i(t)$. The action space is defined as $\mathbb{A} = \{(a, s_1, \ldots, s_m) | a \in [0, S], s_1 \in [0, A_1], \ldots, s_m \in [0, A_m]; \sum_{i=1}^m s_i = a\}$. Where a denotes the total amount of resources used by the controller and s_i is the amount of resources used at station i. Note that although a introduced in the action space is superfluous, it will be useful for writing the cost expressions in a simple manner. The decision epochs are the job arrival times, service completion times and dummy transitions due to uniformization (explained later). When the system is in state $x \in \mathbb{X}$ and action $(a, s) \in \mathbb{A}$ is chosen, the system evolves with the following transition rates

$$q((x, (a, s)), y) = \begin{cases} \lambda_i & \text{if } y = x + e_i, i = 1, \ldots, m, \\ s_i \mu_i & \text{if } y = (x - e_i)^+, i = 1, \ldots, m, , \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

and incurs cost at the rate $C(x(t), (a(t), s(t))) = c(a(t)) + \sum_{i=1}^m h_i x_i(t)$ per unit time at time t.

A decision rule, $d(x)$, maps the state space $\mathbb{X}$ to the action space $\mathbb{A}$. A policy π is a sequence of decision rules. Let Π denote the set of non-anticipating policies, and suppose for any $\pi \in \Pi$ that $\{x^\pi(t), (a^\pi(t), s^\pi(t))); t \geq 0\}$ denotes the stochastic process representing the evolution of state and decisions under policy π. Given the initial state x and discount factor $\alpha > 0$, the $\alpha-$**discounted finite horizon expected cost**, under policy π is given by

$$v_{t,\alpha}^{\pi}(x) := \mathbb{E}_x^{\pi}\left[\int_0^t e^{-\alpha u}[c(a^{\pi}(u)) + \sum_{i=1}^{m} h_i x_i^{\pi}(u)]du\right], \tag{3.3}$$

where $\mathbb{E}_x[\cdot] := \mathbb{E}[\cdot|x(0) = x]$. The **total discounted expected cost** under a policy π given that the initial state of the system is x, is $v_{\alpha}^{\pi}(x) := \lim_{t\to\infty} v_{t,\alpha}^{\pi}(x)$ and the **long-run average cost** is $g^{\pi}(x) := \limsup_{t\to\infty} \frac{v_{t,0}^{\pi}(x)}{t}$. In each case, the optimal costs starting from state x are

$$f^*(x) := \inf_{\pi\in\Pi} f^{\pi}(x), \tag{3.4}$$

for $f = v_{\alpha}$ or g and the policy, π^*, that attains the infimum for all $x \in \mathbb{X}$, is optimal in the respective case. Our objective is to find an allocation policy π that minimizes the total expected discounted cost or the expected average cost rate incurred by the system over an infinite planning horizon.

Since the holding rates under any admissible policy are bounded, we can apply *uniformization* in the spirit of Lippman [36] and consider the discrete time equivalent of the continuous time Markov chain described above. The uniformization rate is chosen to be $\gamma := \sum_{i=1}^{m}(\lambda_i + A_i\mu_i)$. Without loss of generality let $\gamma = 1$. Note that uniformization induces a discount factor $\beta = \frac{\gamma}{\alpha+\gamma}$ for the discrete time chain. When the system is in state $x \in \mathbb{X}$ and action $(a, s) \in \mathbb{A}$ is chosen, the uniformized system evolves with the following transition probabilities

$$p((x,(a,s)),y) = \begin{cases} \lambda_i & \text{if } y = x + e_i, i = 1,\dots,m, \\ s_i\mu_i & \text{if } y = (x-e_i)^+, i = 1,\dots,m, \\ (1 - \sum_{i=1}^{m}(\lambda_i + s_i\mu_i\mathbb{1}_{x_i>0})) & \text{if } y = x, \\ 0 & \text{otherwise} \end{cases}, \tag{3.5}$$

Define $v_{\alpha,k}(x)$ be the minimum total α-discounted expected cost that can be obtained during the last k transitions when starting from state x and let $v_{\alpha,0}(x) = 0$ for all $x \in \mathbb{X}$. Using standard arguments of Markov decision theory [11], the discrete-time finite horizon optimality equations (FHOE) for the system can be written

$$v_{\alpha,k+1}(x) = \min_{(a,s)\in\mathbb{A}} \left\{ \frac{\sum_{i=1}^{m} h_i x_i + c(a)}{\alpha+1} + \beta\left(\sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_i) + \sum_{i=1}^{m} s_i \mu_i v_{\alpha,k}(x-e_i)^+ \right.\right.$$
$$\left.\left. + (1 - \sum_{i=1}^{m} (\lambda_i + s_i \mu_i \mathbb{1}_{x_i>0})) v_{\alpha,k}(x) \right) \right\} \tag{3.6}$$

where $a = \sum_{i=1}^{m} s_i$ is the total resource utilization and e_i denotes the vector of all zeros except one in the i^{th} coordinate. Let $\Delta^i f(n)$ denote the directional first difference in the i^{th} coordinate of a real valued function f defined on $(\mathbb{Z}^+)^m$ so that

$$\Delta^i f(x) = \begin{cases} 0, & \text{for } x_i = 0; \\ f(x) - f(x-e_i), & \text{for } x_i > 0. \end{cases} \tag{3.7}$$

Using the definition in (3.7) and since $\beta = \frac{1}{1+\alpha}$, we can rewrite (3.6) as

$$v_{\alpha,k+1}(x) := \mathcal{T}_\alpha^L v_{\alpha,k}(x) \tag{3.8}$$
$$= \frac{1}{\alpha+1}\left(\sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_i) + \sum_{i=1}^{m} A_i \mu_i v_{\alpha,k}(x) + \min_{(a,s)\in\mathbb{A}} \left\{ c(a) - \sum_{i=1}^{m} s_i \mu_i \Delta^i v_{\alpha,k}(x) \right\} \right) \tag{3.9}$$

where $\mathcal{T}_\alpha^L$ is the dynamic programming mapping defined on real valued functions on $(\mathbb{Z}^+)^m$. The next Lemma shows that the discounted cost value function defined in (3.4) can be obtained by repeatedly applying the mapping $\mathcal{T}_\alpha^L$ starting from an arbitrary function $v_{\alpha,0}$. This is called the value iteration algorithm.

Lemma 3.2.1. *For any discount faction $\alpha > 0$, the following hold,*

(i) the discount cost value function v_α exists and is a unique solution to the discounted cost optimality equations (DCOE) stated below

$$v_\alpha(x) = \mathcal{T}_\alpha^L v_\alpha(x). \tag{3.10}$$

(ii) for arbitrary $v_{\alpha,0}$, the sequence $\{v_{\alpha,k}\}$ defined by the FHOE (3.8) *converges to v_α.*

Proof. See Appendix.

The average cost optimality equations for this system can be written as

$$g + w(x) = \mathcal{T}_0^L w(x) \tag{3.11}$$

$$= \sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i w(x + e_i) + \sum_{i=1}^{m} A_i \mu_i w(x) + \min_{(a,s)\in\mathbb{A}} \left\{ c(a) - \sum_{i=1}^{m} s_i \mu_i \Delta^i w(x) \right\}. \tag{3.12}$$

When a solution (w, g) to the ACOE (3.11) exists, it is such that g is the optimal average cost (independent of the initial state) and the relative value function, $w(x)$, can be interpreted as the minimum expected cost incurred until the next entry into an arbitrary reference state $x_{\text{ref}} \in \mathbb{X}$.

3.2.3 System with full flexibility

In this system the pool of resources are fully shared amongst these stations. A central controller constantly monitors the number of pending jobs of each time present in the system and decides the number of resources $s_i \in [0, S]$ to be allocated to station i. We formulate this problem as a Markov decision process.

Let $x_i \in \mathbb{Z}_+$ denote the total number of jobs at station i (including the job being processed). Define the state space as $x = \{(x_1, \dots, x_m) \in \mathbb{Z}_+^m\}$. When the system is in state $x \in \mathbb{X}$, the central controller can choose an action in the action space $\mathbb{A} = \{(a, (s_1, \dots, s_m)) : a \in [0, S]; \sum_{i=1}^m s_i \le a; s_i \ge 0, i = 1, \dots, m\}$. Where s_i is the amount of resources to be allocated to station i and a is the total amount of resources to be used. The decision epochs are the job arrival, service completion and dummy transition (due to uniformization) times. When the system is in state $x \in \mathbb{X}$ and action $(a, s) \in \mathbb{A}$ is chosen, the system evolves with the following transition rates

$$q((x, (a, s)), y) = \begin{cases} \lambda_i & \text{if } y = x + e_i, i = 1, \dots, m \\ s_i\mu_i & \text{if } y = (x - e_i)^+, i = 1, \dots, m \\ 0 & \text{otherwise} \end{cases}, \tag{3.13}$$

and incurs cost at the rate $C(x, (a, s)) = c(a) + \sum_{i=1}^m h_i x_i$ per unit time. Since the holding rates under any admissible policy are bounded, we can apply *uniformization* in the spirit of Lippman [36] and consider the discrete time equivalent of the continuous time Markov chain described above. The uniformization rate is chosen to be $\gamma := \sum_{i=1}^m (\lambda_i + A_i\mu_i)$. Without loss of generality let $\gamma = 1$. Note that uniformization induces a discount factor $\beta = \frac{\gamma}{\alpha+\gamma}$ for the discrete time chain. When the system is in state $x \in \mathbb{X}$ and action $(a, s) \in \mathbb{A}$ is chosen, the uniformized system evolves with the following transition probabilities

$$p((x, (a, s)), y) = \begin{cases} \lambda_i & \text{if } y = x + e_i, i = 1, \dots, m, \\ s_i\mu_i & \text{if } y = (x - e_i)^+, i = 1, \dots, m, \\ (1 - \sum_{i=1}^m (\lambda_i + s_i\mu_i)) & \text{if } y = x, \\ 0 & \text{otherwise.} \end{cases} \tag{3.14}$$

Let $v_{\alpha,k}(x)$ be the minimum total α-discounted expected cost that can be obtained during the last k transitions when starting from state x. Using standard arguments of Markov decision theory [11], the discrete-time finite horizon optimality equations (FHOE) for the system can be written as

$$v_{\alpha,k+1}(x) = \min_{(a,s)\in\mathbb{A}} \left\{ \frac{\sum_{i=1}^{m} h_i x_i + c(a)}{\alpha+1} + \beta\Bigg(\sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_i) + \sum_{i=1}^{m} s_i \mu_i v_{\alpha,k}(x-e_i)^+ \right.$$
$$\left. + (1 - \sum_{i=1}^{m} \lambda_i - \sum_{i=1}^{m} s_i \mu_i \mathbb{1}_{x_i>0}) v_{\alpha,k}(x) \Bigg) \right\} \tag{3.15}$$

Using the operator Δ^i defined in (3.7), we can rewrite (3.15) as

$$v_{\alpha,k+1}(x) := \mathcal{T}_\alpha^F v_{\alpha,k}(x) \tag{3.16}$$
$$= \frac{1}{\alpha+1}\Bigg[\sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_i) + (1 - \sum_{i=1}^{m} \lambda_i) v_{\alpha,k}(x)$$
$$+ \min_{0\le a\le S} \left\{ c(a) - \max_{\{s_i:\sum_{i=1}^{m} s_i\le a, s_i\ge 0\}} \left\{ \sum_{i=1}^{m} s_i \mu_i \Delta^i v_{\alpha,k}(x) \right\}\right\}\Bigg] \tag{3.17}$$

where $\mathcal{T}_\alpha^F$ is the dynamic programming mapping defined on real valued functions on $\mathbb{X}$. As in the previous section, we may apply Theorem 6.10.4 of [46] to get $v_{\alpha,k} \uparrow v_\alpha$. Moreover, v_α satisfies the *discounted cost optimality equations* (DCOE) stated below

$$v_\alpha(x) = \mathcal{T}_\alpha^F v_\alpha(x). \tag{3.18}$$

For the long-run average cost problem, let $w(x)$ denote the relative value function for this problem and g denote the optimal average cost rate when they exist and in this case they satisfy the following average cost optimality equations

$$g + w(x) = \mathcal{T}^F w(x). \tag{3.19}$$

3.3 Structural Properties of Optimal Policies

In this section, we discuss the structure of the optimal policy for the limited flexibility and the fully flexible models under both discounted cost and optimal cost criteria. We start be making some connections between the discounted cost value function and relative value function.

Proposition 3.3.1. *Suppose* $\sum_{i=1}^{m} \frac{\lambda_i}{\mu_i} < S$ *and for limited flexibility case* $\frac{\lambda_i}{\mu_i} < A_i, \quad i = 1, 2, \ldots, m$ *and* $\sum_{i=1}^{m} A_i = S$*. The value function in each case satisfies the following properties:*

1. *There exists a finite constant* $g := \lim_{\alpha \uparrow 1}(1-\alpha)v_\alpha(x)$ *for all* $x \in \mathbb{X}$*.*
2. $w_\alpha(x) = v_\alpha(x) - v_\alpha(\mathbf{0})$ *converges along a subsequence to a function* w *on* $\mathbb{X}$ *such that* (g, w) *satisfy the average cost optimality equation.*
3. *Any limit point of a sequence of discounted cost optimal policies (as* $\alpha \uparrow 1$*) is average cost optimal.*

Proof. It is enough to show the existence of a policy with finite average cost that induces a Markov chain that is either irreducible or consists of a single recurrent class with finite absorption times when starting from any of the transient states. The result then follows from Theorems 7.2.3 and 7.5.6 of Sennott [49]. For both limited and full flexibility cases consider a policy that assigns A_i resources at station i at all times. Then the m stations behave as independent $M/M/1$

queues. Based on our assumption, the utilization for station i is $\frac{\lambda_i}{A_i\mu_i} < 1$. Thus the Markov chain induced under this policy is irreducible and positive recurrent. The long-run average number of jobs at station i is $Q_i(\infty) = \frac{\lambda_i}{A_i\mu_i - \lambda_i} < \infty$ and the long-run average cost incurred under this policy is $c(S) + \sum_{i=1}^{m} h_i Q_i(\infty) < \infty$.
■

3.3.1 System with limited flexibility

Our next result shows that it is optimal to prioritize the stations based on $h_i\mu_i$ indices. Without loss of generality assume in the rest of this section that the stations are indexed such that $h_1\mu_1 \geq h_2\mu_2 \geq \ldots \geq h_m\mu_m$. We say that a policy follows the $c - \mu$ rule, if under this policy a higher index station with pending jobs is assigned resources only after all the lower index stations with pending jobs have reached their respective resource capacity limit. i.e,

$$s_k(x) > 0 \iff s_{k'}(x) = A_{k'} \mathbb{1}_{\{x'_k > 0\}} \quad \forall \quad k' = 1, \ldots, k-1.$$

The main result of this section is summarized in the following theorem:

Theorem 3.3.2. *Under the infinite horizon discounted cost criteria or the average cost criteria, the following holds*

1. *(Scheduling) there exists an optimal policy that follows the $c-\mu$ rule i.e, the policy gives highest priority to station 1 then to station 2 and so on with lowest priority to station m.*
2. *(Utilization) for the problem with only two stations ($m = 2$), there exists an optimal policy that uses a total of $a(i, j)$ resources in state $x = (i, j)$ such that $a(i, j)$ is increasing in both i and j.*

We next prove (1) of Theorem 3.3.2. For this we start by assuming that the utilization cost rate is zero ($c(u) = 0$). We will relax this assumption later. The main idea is that by adopting a $c - \mu$ scheduling policy, the long-run discounted cost or the average cost under is minimized for any given resource utilization policy. Let $V_k^\pi(t) \in \mathbb{R}^+$ denote the total remaining service time of jobs (workload) at station k under policy π at time t. We claim that the $c - \mu$ priority policy minimizes path-wise the criteria $\sum_{k=1}^m r_k V_k(t)$ for all $t \geq 0$, where $r_1 \geq r_2 \geq \cdots \geq r_m \geq 0$. Our proof is analogous to [41]. We use the following Lemma of [41] to show this result.

Lemma 3.3.3. *(Nain '94) Let $(a_1, a_2, \ldots, a_m)$ and $(b_1, b_2, \ldots, b_m)$ be vector in $(\mathbb{R}^+)^m$ s.t. $\sum_{k=1}^{\ell} a_k \leq \sum_{k=1}^{\ell} b_k$ for $\ell = 1, \ldots, m$ then,*

$$\sum_{k=1}^{m} r_k a_k \leq \sum_{k=1}^{m} r_k b_k.$$

Proposition 3.3.4. *The $c - \mu$ priority policy minimizes the criteria $\sum_{k=1}^m r_k V_k^\pi(t)$ for all $t \geq 0$ path-wise, where $r_1 \geq r_2 \geq \cdots \geq r_m \geq 0$.*

Proof. Without loss of generality assume that the stations are empty at $t = 0^-$. Let $\pi \in \Pi$ be any non-anticipating policy and $a^\pi(t) = \sum_{k=1}^m s_k^\pi(t)$ be the total resource utilization under this policy at time t. Let γ be a policy that uses the $c - \mu$ priority rule to decide how much resources to use at each station but has the same (possibly less) total resource utilization as the policy π at any time ($a^\pi(t)$). Consider two processes on the same probability space. So that both processes see a same sequence of arrivals and departures. Process 1(2) follows policy $\pi(\gamma)$. Let $A_{n,k}$ denote the arrival time of the n^{th} job at station k, $D_{n,k}^\pi$ the departure time of the n^{th} job at station k under policy π and $D_{n,k}^\gamma$ the departure time of the n^{th} job at station k under policy γ. Let $\{t_n\}_{n=1}^\infty$, $t_1 \leq t_2 \leq \ldots$ be the sequence of event times obtained by the superposition of the processes $\{A_{n,k}\}$, $\{D^{\pi}_{n,k}\}$ and $\{D^{\gamma}_{n,k}\}$. Thus, t_n denotes the time of n^{th} event when the three processes are combined. Note that since the arrival processes are Poisson, $\lim_{n\to\infty} A_{n,k} = +\infty$ almost surely. Therefore $\lim_{n\to\infty} t_n = +\infty$ a.s.. We will show that

$$\sum_{k=1}^{\ell} V_k^{\gamma}(t) \le \sum_{k=1}^{\ell} V_k^{\pi}(t) \quad \ell = 1, 2, \ldots, m \quad (a.s.) \tag{3.20}$$

for all $t \ge 0$ using induction on the sequence of event times $\{t_n\}_{n=1}^{\infty}$.

Consider $\omega \in \Omega$. For the basis step, note that since the stations are empty at the start we have,

$$\sum_{k=1}^{\ell} V_k^{\gamma}(t,\omega) = \sum_{k=1}^{\ell} V_k^{\pi}(t,\omega) \quad \ell = 1, 2, \ldots, m; \quad 0 \le t \le t_1$$

For the induction step, assume that (3.20) holds for all times up to the time of n^{th} event in the superposed process i.e,

$$\sum_{k=1}^{\ell} V_k^{\gamma}(t,\omega) \le \sum_{k=1}^{\ell} V_k^{\pi}(t,\omega) \quad \ell = 1, 2, \ldots, m; \quad 0 \le t \le t_n.$$

Consider the time interval $t_n < t < t_{n+1}$. Let $(s_1^{\gamma}(t,\omega), s_2^{\gamma}(t,\omega), \ldots, s_m^{\gamma}(t,\omega))$ be the resource utilization decisions made by the policy γ during this time interval. The workload is

$$V_k^{\gamma}(t,\omega) = \begin{cases} 0 & \text{if } V_k^{\gamma}(t_n,\omega) = 0 \\ V_k^{\gamma}(t_n,\omega) - \int_{t_n}^{t} s_k^{\gamma}(y,\omega)dy & \text{if } V_k^{\gamma}(t_n,\omega) > 0 \end{cases} \quad t_n < t < t_{n+1}.$$

Since the policy γ uses $c-\mu$ priority rule to assign resources to various stations, we have

$$\sum_{k=1}^{\ell} \int_{t_n}^{t} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} s_k^{\gamma}(y,\omega)dy \ge \sum_{k=1}^{\ell} \int_{t_n}^{t} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} s_k^{\pi}(y,\omega)dy \quad \ell = 1, 2, \ldots, m; \quad t_n < t < t_{n+1}. \tag{3.21}$$

Therefore for $\ell = 1, 2, \ldots, m$ and $t_n < t < t_{n+1}$ we have,

$$\sum_{k=1}^{\ell} V_k^{\gamma}(t,\omega) = \sum_{k=1}^{\ell} V_k^{\gamma}(t,\omega)\mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)=0\}} + \sum_{k=1}^{\ell} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}}\Big(V_k^{\gamma}(t_n,\omega) - \int_{t_n}^{t} s_k^{\gamma}(y,\omega)dy\Big), \tag{3.22}$$

$$\begin{aligned}
&\leq \sum_{k=1}^{\ell} V_k^{\pi}(t,\omega)\mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)=0\}} + \sum_{k=1}^{\ell} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} V_k^{\pi}(t_n,\omega) - \sum_{k=1}^{\ell} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} \int_{t_n}^{t} s_k^{\gamma}(y,\omega)dy, \\
&\leq \sum_{k=1}^{\ell} V_k^{\pi}(t,\omega)\mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)=0\}} + \sum_{k=1}^{\ell} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} V_k^{\pi}(t_n,\omega) - \sum_{k=1}^{\ell} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} \int_{t_n}^{t} s_k^{\pi}(y,\omega)dy, \\
&= \sum_{k=1}^{\ell} V_k^{\pi}(t,\omega)\mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)=0\}} + \sum_{k=1}^{\ell} \mathbb{1}_{\{V_k^{\gamma}(t_n,\omega)>0\}} V_k^{\pi}(t,\omega) = \sum_{k=1}^{\ell} V_k^{\pi}(t,\omega),
\end{aligned}$$

where the second inequality follows by the induction hypothesis and the third inequality follows by (3.21).

At $t = t_{n+1}$, if the event that occurs is an arrival event at some station, then the workload under both π and γ increases by the same amount at that station so the result still holds by (3.22) at $t = t_{n+1}$. If the event is a departure in either π or γ, then

$$V_k^{\pi}(t_{n+1},\omega) = V_k^{\pi}(t_{n+1}^{-},\omega), \quad V_k^{\gamma}(t_{n+1},\omega) = V_k^{\gamma}(t_{n+1}^{-},\omega).$$

Thus the result again holds by (3.22) at $t = t_{n+1}$. Now the main result of this proposition follows from the Lemma 3.3.3. ■

We next show that the $c - \mu$ priority policy minimizes the criteria $\sum_{k=1}^{m} h_k\, \mathbb{E}[x_k(t)]$ for all $t \geq 0$, where $x_k(t)$ denotes the number of jobs at station k at time t.

Lemma 3.3.5. *For any policy $\pi \in \Pi$ we have,*

$$\mathbb{E}[V_k^{\pi}(t)] = \mathbb{E}[x_k^{\pi}(t)]/\mu_k.$$

Proof. Let $\sigma^{\pi}_{j,k}(t)$ denote the service time of the j^{th} customer waiting at station k at time t. Then,

$$V^{\pi}_{k}(t) = \sum_{j=1}^{x^{\pi}_{k}(t)} \sigma^{\pi}_{j,k}(t).$$

Since the state only includes the queue lengths and the policy is non-anticipating, π does not know about the present and future service times. Furthermore, by the memoryless property of exponentially distributed service times we have,

$$\mathbb{E}[V^{\pi}_{k}(t)] = \mathbb{E}[\sum_{j=1}^{x^{\pi}_{k}(t)} \sigma^{\pi}_{j,k}(t)] = \mathbb{E}[\mathbb{E}[\sum_{j=1}^{x^{\pi}_{k}(t)} \sigma^{\pi}_{j,k}(t)|x^{\pi}_{k}]] = \mathbb{E}[x^{\pi}_{k}(t)]/\mu_k.$$

■

Considering $r_k = h_k\mu_k$, the results of Proposition 3.3.4 and Lemma 3.3.5 shows that the $c-\mu$ priority policy minimizes the criteria $\sum_{k=1}^{m} h_k\, \mathbb{E}[x_k(t)]$ for all $t \geq 0$.

Note that in the proof of Proposition 3.3.4, we constructed a $c-\mu$ priority policy γ that uses the same (possibly less) total amount of resources as the original policy π yet incurred less holding costs. So even if the utilization cost rate, ($c(\cdot)$), is not zero, the $c-\mu$ priority policy still minimizes the criteria $\sum_{k=1}^{m} h_k\, \mathbb{E}[x_k(t)]$ for all $t \geq 0$. Thus the $c-\mu$ policy is optimal for the discounted cost and the average cost criteria (assuming stability). This proves Statement (1) of Theorem 3.3.2.

Next we provide a characterization of how the resource utilization at individual stations ($s^{\pi}_{k}(t)$) depends on the queue lengths. To this end we will establish several properties that the discounted cost value function using the DCOE

(3.8) introduced in Section 2.2. Recall that the FHOE are:

$$
\begin{aligned}
v_\alpha^k(x) &:= T_\alpha^L v_\alpha^k(x) \\
&= \sum_{i=1}^m h_i x_i + \sum_{i=1}^m \lambda_i v_\alpha^k(x+e_i) + \sum_{i=1}^m A_i \mu_i v_\alpha^k(x) + \min_{s \in \mathbb{A}} \Big\{ c(a) - \sum_{i=1}^m s_i \mu_i \Delta^i v_\alpha^k(x) \Big\} \Big].
\end{aligned}
$$

We start with the following intuitive lemma.

Lemma 3.3.6. *(Monotonicity) The optimal discounted expected cost value function $v_\alpha(x)$ is non-decreasing in x. That is, $v_\alpha(x+e_j) \geq v_\alpha(x)$, $j = 1, \ldots, m$.*

3.3.2 Two Stations

For the remainder of this section, assume that the number of stations $m = 2$. To simplify exposition, we will also denote the state variable $x = (i, j)$.

Lemma 3.3.7. *The discounted cost value function satisfies the following properties,*

1. *(convexity)*
   - *(a)* $\Delta^1 v_\alpha(i+1, j) \geq \Delta^1 v_\alpha(i, j)$ *for* $i \geq 1, j \geq 0$
   - *(b)* $\Delta^2 v_\alpha(i, j+1) \geq \Delta^2 v_\alpha(i, j)$ *for* $i \geq 0, j \geq 1$
2. *(super-modularity)* $\Delta^1 v_\alpha(i, j+1) \geq \Delta^1 v_\alpha(i, j)$ *for* $i \geq 1, j \geq 0$
3. - *(a)* $\Delta^1 v_\alpha(i+1, j) \geq \Delta^1 v_\alpha(i, j+1)$ *for* $i \geq 1, j \geq 0$
   - *(b)* $\Delta^2 v_\alpha(i, j+1) \geq \Delta^2 v_\alpha(i+1, j)$ *for* $i \geq 0, j \geq 1$

Proof. See Appendix.

Given the value function, the DCOE can be used to obtain the optimal amount of resources to use at any station by solving the following optimization problem:

$$\max_{s\in\mathbb{A}}\Big\{\sum_{i=1}^{2} s_i\mu_i\Delta^i v_{\alpha,k}(x) - c(\sum_{i=1}^{2} s_i)\Big\}. \tag{3.23}$$

For $y \in \mathbb{R}^2_+$, define the function:

$$\psi(y) = \operatorname{argmax}_{s\in\mathbb{A}}\Big\{y^T s - c(\sum_{i=1}^{2} s_i)\Big\}, \tag{3.24}$$

where the arg max is the smallest s^* that attains the max. By the assumptions on $c(\cdot)$, the max in (3.24) is attained by some $s^* \in \mathbb{A}$ for all $y \geq 0$. For any state x, let $y(x) = (\mu_1\Delta^1 v_\alpha(x), \mu_2\Delta^2 v_\alpha(x))$ then the optimal resource allocation in state x is $s^* = \psi(y(x))$. For comparison of vectors, let $(\geq)$ be the usual coordinate-wise order:

$$(x_1, \ldots, x_m) \geq (y_1, \ldots, y_m) \iff x_i \geq y_i \quad \forall i.$$

The following lemma says that the total resource utilization is non-decreasing in y. This is a direct consequence of the fact that the function $c(\cdot)$ is convex and non-decreasing.

Lemma 3.3.8. *Let $y, y' \in \mathbb{R}^2_+$ such that $y \geq y'$ (coordinate-wise order) then $\sum_{i=1}^2 \psi_i(y) \geq \sum_{i=1}^2 \psi_i(y')$.*

Proof. For ease of exposition, consider $c(\cdot)$ to be continuously differentiable and strictly convex so that the function $c'^{-1}(\cdot)$ is well defined. It can be readily verified that (3.23) is a convex optimization problem, using KKT conditions (see

Boyd [12]) the solution can be written in closed form as:

$$(\psi_1(y), \psi_2(y)) = \begin{cases} (c'^{-1}(y_1), 0) & \text{if } c'(A_1) \geq y_1 \geq y_2 \\ (A_1, 0) & \text{if } y_1 \geq c'(A_1) \geq y_2 \\ (A_1, c'^{-1}(y_2) - A_1) & \text{if } y_1 \geq y_2; c'(A_1 + A_2) \geq y_2 \geq c'(A_1) \\ (A_1, A_2) & \text{if } y_1 \geq y_2 \geq c'(A_1 + A_2) \\ (0, c'^{-1}(y_2)) & \text{if } c'(A_2) \geq y_2 \geq y_1 \\ (0, A_2) & \text{if } y_2 \geq c'(A_2) \geq y_1 \\ (c'^{-1}(y_1) - A_2, A_2) & \text{if } y_2 \geq y_1; c'(A_1 + A_2) \geq y_1 \geq c'(A_2) \\ (A_1, A_2) & \text{if } y_2 \geq y_1 \geq c'(A_1 + A_2) \end{cases} .$$

Result follows since $c(\cdot)$ is convex and non-decreasing. ■

Proposition 3.3.9. *Under the infinite horizon discounted cost criteria or the average cost criteria, there exists an optimal policy in which the total amount of resources used by all stations is increasing in both i and j.*

Proof. Properties (1) and (2) of Lemma 3.3.7 states that $\Delta^\ell v_\alpha(i, j)$ is non-decreasing in both i and j for $\ell = 1, 2$. Thus by Lemma 3.3.8, the total amount of resources used are non-decreasing in both i and j.

As an immediate consequence of Proposition 3.3.1, the relative value function w in the ACOE satisfies the properties of Lemma 3.3.7 with v_α replaced by w. Therefore, result holds for the average cost criteria too. ■

3.3.3 System with full flexibility

We start this section by stating the main theorem that characterizes the optimal policy for the fully flexible case. We continue to assume that the stations are indexed such that $h_1\mu_1 \geq h_2\mu_2 \geq \ldots \geq h_m\mu_m$.

Theorem 3.3.10. *Under the infinite horizon discounted cost criteria or the average cost criteria, the following holds*

1. *there exists an optimal policy that at any time serves exactly one station.*
2. *(scheduling) there exists an optimal policy that follows the $c-\mu$ rule i.e, the policy gives highest priority (serves exhaustively) to station 1 then to station 2 and so on with lowest priority to station m.*
3. *(resource utilization) for the problem with only two stations ($m = 2$), if $\mu_1 = \mu_2$, then there exists an optimal policy that uses $a(i, j)$ resources in state (i, j) such that $a(i, j)$ is increasing in both i and j.*

Note that based on the system state x, the decision maker needs to answer two questions - how many resources to use (utilization) and how to distribute these resources among various stations (scheduling). Results 1 and 2 of Theorem 3.3.10 addresses the issue of scheduling and result 3 of Theorem 3.3.10 shows that for the particular case of two stations with $\mu_1 = \mu_2$, the resource utilization is increasing in the length of queues. The rest of the section is devoted to proving Theorem 3.3.10. Proving these results requires us to show several structural properties that the discounted cost value function satisfies. The next lemma states the intuitive result that the value function is increasing in each component of x.

Lemma 3.3.11. *(Monotonicity) The optimal discounted expected cost value function $v_\alpha(x)$ is increasing in x. That is, $v_\alpha(x+e_j) \geq v_\alpha(x)$, $j = 1, \ldots, m$.*

Proof. We use induction and FHOE (3.15) to prove the result. Let $v_{\alpha,0}(x) = 0$, the result holds trivially for $k = 0$. For the inductive step, suppose $v_{\alpha,k}(x)$ satisfies the result for all $x \in \mathbb{X}$. Suppose (a, s) is the optimal decision for $k+1$ stage problem when the initial state is $x + e_j$. The FHOE (3.15) yields,

$$\begin{aligned}
(\alpha+1)v_{\alpha,k+1}(x+e_j) &= c(a) + \sum_{i=1}^{m} h_i x_i + h_j + \sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_j+e_i) + \sum_{i=1}^{m} s_i \mu_i v_{\alpha,k}((x+e_j-e_i)^+) \\
&\quad +(1 - \sum_{i=1}^{m} \lambda_i - \sum_{i=1}^{m} s_i\mu_i) v_{\alpha,k}(x+e_j) \\
&\geq c(a) + \sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_i) + \sum_{i=1}^{m} s_i \mu_i v_{\alpha,k+1}((x-e_i)^+) \\
&\quad +(1 - \sum_{i=1}^{m} \lambda_i - \sum_{i=1}^{m} s_i\mu_i) v_{\alpha,k}(x) \\
&\geq (\alpha+1) v_{\alpha,k+1}(x), \qquad (3.25)
\end{aligned}$$

where, the first inequality follows from the induction hypothesis and the second inequality follows since (a, s) is potentially sub-optimal for the state x. The result follows since $v_{\alpha,k}(x)) \uparrow v_\alpha(x)$ for all $x \in \mathbb{X}$. ∎

Proposition 3.3.12. *There exists an optimal policy that at any time serves exactly one station.*

Proof. The DCOE (3.18) implies that the optimal allocation decision can be obtained by solving the following optimization problem:

$$\min_{0 \leq a \leq S} \left\{ c(a) - \max_{\{s_i : s_i \geq 0, \sum_{i=1}^{m} s_i \leq a\}} \left\{ \sum_{i=1}^{m} s_i \mu_i \Delta^i v_\alpha(x) \right\} \right\}. \qquad (3.26)$$

Suppose a^* is the optimal capacity to use in state x. The decision about how a^* will be split amongst various stations is obtained by solving the optimization problem:

$$\max_{\{s_i : s_i \geq 0, \sum_{i=1}^{m} s_i \leq a^*\}} \left\{ \sum_{i=1}^{m} s_i \mu_i \Delta^i v_\alpha(x) \right\}. \tag{3.27}$$

Since the objective function is linear in s, the optimal solution is an extreme point of the feasible set of solutions $\{s_i : s_i \geq 0, \sum_{i=1}^{m} s_i \leq a^*\}$ and the optimal allocation decision in this has the form $s^* = a^* e_i$ for some i. The solution can be made more explicit by letting i^* be the index corresponding to the station with the largest $\mu_i \Delta^i v_\alpha(x)$ index. Proposition 3.3.11 shows that $\Delta^i v_\alpha(x) \geq 0$, and the objective function in (3.26) that the optimal policy is to serve only at station i^*. ∎

Proposition 3.3.12 showed the existence of an optimal policy that works at exactly one station at any given time. Denote the set of all such policies by $\Pi_R \subseteq \Pi$. Recall that the stations are indexed such that $h_1\mu_1 \geq h_2\mu_2 \geq \ldots \geq h_m\mu_m$. We next show that there exists an optimal policy that selects the station to work on based on a index rule such that the highest priority is given to station 1 then to station 2 and so on with lowest priority to station m. This is the well known $c - \mu$ rule.

For this we start by assuming that the utilization cost rate is zero ($c(u) = 0$). We will relax this assumption later. Let $V_k^\pi(t) \in \mathbb{R}^+$ denote the total remaining service time of jobs (workload) at station k under policy π at time t. We first show that the $c - \mu$ priority policy minimizes path-wise the criteria $\sum_{k=1}^{m} r_k V_k(t)$ for all $t \geq 0$, where $r_1 \geq r_2 \geq \cdots \geq r_m \geq 0$. Our proof is analogous to [41].

Proposition 3.3.13. *The $c - \mu$ priority policy minimizes the cost $\sum_{k=1}^{m} r_k V_k^\pi(t)$ for all*

$t \geq 0$ path-wise, where $r_1 \geq r_2 \geq \cdots \geq r_m \geq 0$.

Proof. Without loss of generality assume that the stations are empty at $t = 0^-$. Let $\pi \in \Pi^R$ be any non-anticipating policy and $a^\pi(t)$ be the total resource utilization under this policy at time t. Let γ be a policy that uses the $c - \mu$ preemptive priority rule to decide at which station to work at but has the same (possibly less) total resource utilization as the policy π at any time ($a^\pi(t)$). Consider two processes on the same probability space. Process 1(2) follows policy $\pi(\gamma)$ and process 2 follows policy γ. Let $A_{n,k}$ denote the arrival time of the n^{th} job at station k, $D^\pi_{n,k}$ the departure time of the n^{th} job at station k under policy π and $D^\gamma_{n,k}$ the departure time of the n^{th} job at station k under policy γ. Let $\{t_n\}_{n=1}^\infty$, $t_1 \leq t_2 \leq \ldots$ be the sequence of event times obtained by the superposition of the processes $\{A_{n,k}\}$, $\{D^\pi_{n,k}\}$ and $\{D^\gamma_{n,k}\}$. Note that since the arrival processes are Poisson, $\lim_{n\to\infty} A_{n,k} = +\infty$ almost surely. Therefore $\lim_{n\to\infty} t_n = +\infty$ a.s.. We show that

$$\sum_{k=1}^{\ell} V_k^\gamma(t) \leq \sum_{k=1}^{\ell} V_k^\pi(t) \quad \ell = 1, 2, \ldots, m \quad (a.s.) \tag{3.28}$$

for all $t \geq 0$ using induction on the sequence of event times $\{t_n\}_{n=1}^\infty$.

Consider $\omega \in \Omega$. For the basis step, note that since the stations are empty at the start we have,

$$\sum_{k=1}^{\ell} V_k^\gamma(t,\omega) = \sum_{k=1}^{\ell} V_k^\pi(t,\omega) = 0 \quad \ell = 1, 2, \ldots, m; \quad 0 \leq t \leq t_1$$

For the induction step, assume that (3.20) holds for all times upto the time of n^{th} event in the superposed process i.e,

$$\sum_{k=1}^{\ell} V_k^\gamma(t,\omega) \leq \sum_{k=1}^{\ell} V_k^\pi(t,\omega) \quad \ell = 1, 2, \ldots, m; \quad 0 \leq t \leq t_n.$$

We will show that (3.28) holds for $t_n < t \leq t_{n+1}$. First, consider the time interval $t_n < t < t_{n+1}$. Let $a^\gamma(t,\omega)$ be the resource utilization decision made by the policy γ during this time interval. If $\sum_{k=1}^{\ell} V_k^\gamma(t_n,\omega) = 0$, then (3.28) holds. Consider the case when $\sum_{k=1}^{\ell} V_k^\gamma(t_n,\omega) > 0$, and let $i = \min\{k|V_k^\gamma(t_n,\omega) > 0\}$. So i is the lowest indexed station with some pending work at time t_n. Since γ uses the $c-\mu$ priority rule, the evolution of workload for $t_n < t \leq t_{n+1}$ is

$$V_k^\gamma(t,\omega) = \begin{cases} 0 & \text{if } k < i \\ V_k^\gamma(t_n,\omega) - \int_{t_n}^{t} a^\gamma(s,\omega)ds & \text{if } k = i \\ V_k^\gamma(t_n,\omega) & \text{if } k > i. \end{cases}$$

Therefore for $\ell = 1,2,\ldots,i-1$ and $t_n < t < t_{n+1}$ we have,

$$\sum_{k=1}^{\ell} V_k^\gamma(t,\omega) = 0 \leq \sum_{k=1}^{\ell} V_k^\pi(t,\omega), \tag{3.29}$$

and for $\ell \geq i$,

$$\begin{aligned} \sum_{k=1}^{\ell} V_k^\gamma(t,\omega) &= \sum_{k=1}^{\ell} V_k^\gamma(t_n,\omega) - \int_{t_n}^{t} a^\gamma(s,\omega)ds, \\ &\leq \sum_{k=1}^{\ell} V_k^\pi(t_n,\omega) - \int_{t_n}^{t} a^\pi(s,\omega)ds, \\ &\leq \sum_{k=1}^{\ell} V_k^\pi(t,\omega), \end{aligned} \tag{3.30}$$

where the second inequality follows by the induction hypothesis and since $a^\gamma(s,\omega) = a^\pi(s,\omega)$ by construction.

Now at $t = t_{n+1}$ if the event that occurs is an arrival event at some station, then the workload under both π and γ increases by the same amount at that station so the result still holds by (3.22) at $t = t_{n+1}$. If the event is a departure in either π or γ, then

$$V_k^{\pi}(t_{n+1}, \omega) = V_k^{\pi}(t_{n+1}^{-}, \omega), \quad V_k^{\gamma}(t_{n+1}, \omega) = V_k^{\gamma}(t_{n+1}^{-}, \omega).$$

Thus the result again holds by (3.29) and (3.30) at $t = t_{n+1}$. Now the main result of this proposition follows from the Lemma 3.3.3. ∎

Recall that $x_k(t)$ denotes the number of jobs at station k at time t. Considering $r_k = h_k\mu_k$, the results of Proposition 3.3.13 and Lemma 3.3.5 shows that the $c-\mu$ priority policy minimizes the criteria $\sum_{k=1}^{m} h_k \,\mathbb{E}[x_k(t)]$ for all $t \geq 0$. Note that in the proof of Proposition 3.3.4, we constructed a $c-\mu$ priority policy γ that uses the same (possibly less) total amount of resources as the original policy π. Thus even if the utilization cost rate, $(c(\cdot))$, is not zero, the $c-\mu$ priority policy is optimal. Thus the $c-\mu$ policy is optimal for the discounted cost and the average cost criteria (assuming stability). This proves Statement (2) of Theorem 3.3.10.

Properties 1 and 2 of Theorem 3.3.10 completely characterize the scheduling policy i.e, how resources are to be allocated at different stations. We can use these results to further simplify the dynamic programming mapping (3.16). For any state x, suppose i^* be the index of highest priority queue with at least one pending job. Since the stations are indexed such that $h_1\mu_1 \geq h_2\mu_2 \geq \ldots h_m\mu_m$, i^* is the least indexed station for which $x_{i^*} \geq 1$. The dynamic programming mapping (3.16) can be written

$$\mathcal{T}_{\alpha}^{F} v_{\alpha}(x) = \frac{1}{\alpha+1}\Big[\sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i v_{\alpha}(x+e_i) + \min_{0\leq a\leq S}\Big\{c(a) + (1 - \sum_{i=1}^{m}\lambda_i - \mu_{i^*}a)v_{\alpha}(x) - \mu_{i^*}v_{\alpha}(x-e_{i^*}\Big\}\Big]. \tag{3.31}$$

Next we provide a characterization of how the resource utilization under the optimal policy depends on the queue lengths (3 of Theorem 3.3.10). For the remainder of this section, assume that the number of stations, $m = 2$ and $\mu_1 = \mu_2 = \mu$.

We show several properties that the discounted cost value function satisfies in the following lemma.

Lemma 3.3.14. *If $h_1 \geq h_2$, then the following holds,*

1. *(diagonal dominance) $v_\alpha(i+1, j-1) \geq v_\alpha(i, j)$ for $i \geq 0, j \geq 1$.*
2. *$\Delta^1 v_\alpha(i, j) \geq \Delta^2 v_\alpha(i, j)$ for $i \geq 1, j \geq 1$,*
3. *(convexity)*
   (a) *$\Delta^1 v_\alpha(i+1, j) \geq \Delta^1 v_\alpha(i, j)$ for $i \geq 1, j \geq 0$,*
   (b) *$\Delta^2 v_\alpha(i, j+1) \geq \Delta^2 v_\alpha(i, j)$ for $i \geq 0, j \geq 1$,*
4. *(super-modularity) $\Delta^1 v_\alpha(i, j+1) \geq \Delta^1 v_\alpha(i, j)$ for $i \geq 1, j \geq 0$,*
5. (a) *$\Delta^1 v_\alpha(i+1, j) \geq \Delta^1 v_\alpha(i, j+1)$ for $i \geq 1, j \geq 0$,*
   (b) *$\Delta^2 v_\alpha(i, j+1) \geq \Delta^2 v_\alpha(i+1, j)$ for $i \geq 0, j \geq 1$.*

Proof. See Appendix.

To derive structural results for the optimal policy, we use the following definitions (see George and Harrison[23]) to simplify the optimality equations further. Let $\mathbb{A} = [0, S]$ and define

$$\phi(y) = \sup_{a \in \mathbb{A}} \{ay - c(a)\}, \tag{3.32}$$

$$\psi(y) = \arg\max_{a \in \mathbb{A}} \{ay - c(a)\},$$

where the arg max is the smallest a^* that attains the supremum. By the assumptions on $c(\cdot)$ and A, the supremum in (3.32) is finite for all $y \geq 0$ and it is attained by some $a^* \in \mathbb{A}$. Using these definitions the dynamic programming mapping (3.31) can be written in simplified form as

$$\mathcal{T}_{\alpha}^{F} v_{\alpha}(x) = \frac{1}{\alpha+1}\Big[\sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i v_{\alpha}(x+e_i) + \min_{a\in\mathbb{A}}\Big\{c(a) + (1-\sum_{i=1}^{m}\lambda_i - \mu_{i^*}a)v_{\alpha}(x) + \mu_{i^*}v_{\alpha}(x-e_{i^*}\Big\}\Big] \tag{3.33}$$

$$= \frac{1}{\alpha+1}\Big[\sum_{i=1}^{m} h_i x_i + \sum_{i=1}^{m} \lambda_i v_{\alpha}(x+e_i) + (1-\sum_{i=1}^{m}\lambda_i)v_{\alpha}(x) - \phi(\mu_{i^*}\Delta^{i^*}v_{\alpha}(x))\Big], \tag{3.34}$$

where i^* is the index of the highest priority station with pending jobs i.e, the least indexed station for which $x_{i^*} \geq 1$.

In order to derive structural results for an optimal discounted expected cost policy, we make use of several important properties of functions $\phi(y) = \max_{a\in\mathbb{A}}\{ya - c(a)\}$ and its associated maximizers $\psi(y) = \arg\max_{a\in\mathbb{A}}\{ya - c(a)\}$ that were introduced in the mapping $\mathcal{T}_{\alpha}^{F}$ (3.34). Recall that the *conjugate* of $c(\cdot)$, $\phi(\cdot)$, is convex (cf. [12]). Moreover, $\psi(y)$ is left continuous, non-decreasing and equals $\phi'(y)$ wherever the derivative exists. Furthermore, neither $\phi(\cdot)$ of $\psi(\cdot)$ get affected if the utilization cost function $c(\cdot)$ is replaced by its convex hull $\hat{c}(\cdot)$ (cf. Section 4 of [23]). Where the convex hull is defined to be the largest convex function f on A such that $c(x) \geq f(x)$ for all $x \in A$. This allows us to model a discrete capacity set $\hat{A} = \{s_1, s_2, \ldots, s_p\}$ by considering a piecewise linear cost rate function $\hat{c}(\cdot)$. In particular, consider the utilization cost rate function to be the convex hull of the associated utilization cost rates $c(s_i)$, which is piece-wise linear, along with a connected set $A = [s_1, s_p]$, then the optimal solution always comes from $\hat{A}$.

The next proposition proves Property 3 of Theorem 3.3.10 and shows that the resources to be used are increasing in the queue lengths of both stations.

Proposition 3.3.15. *Under the infinite horizon discounted cost criteria and the average cost criteria, for the problem with only two stations* ($m = 2$), *if* $\mu_1 = \mu_2$, *then there exists an optimal policy that uses* $a(i, j)$ *resources in state* (i, j) *such that* $a(i, j)$ *is increasing in both* i *and* j.

Proof. From (3.34), the optimal amount of resources to use in state (i, j) is,

$$a(i, j) = \begin{cases} \psi(\mu\Delta^1 v_\alpha(i, j)) & \text{if } i \geq 1, j \geq 0 \\ \psi(\mu\Delta^2 v_\alpha(i, j)) & \text{if } i = 0, j \geq 1. \end{cases}$$

The result follows since $\psi(\cdot)$ is non-decreasing and the value function satisfies properties 2, 3 (a), (b) and 4 of Lemma 3.3.14.

As an immediate consequence of Proposition 3.3.1, the relative value function w in the ACOE satisfies the properties of Lemma 3.3.14 with v_α replaced by w. Therefore, the result holds for the average cost criteria. ∎

3.4 Comparison of Systems - Numerical Study

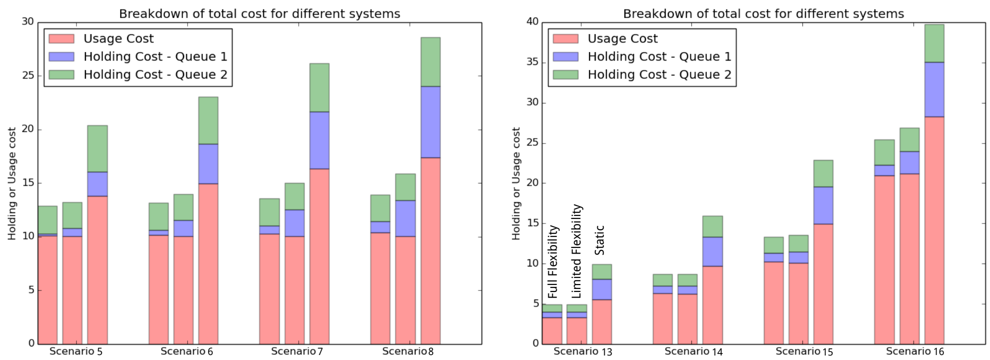
In this section we consider several numerical examples and compare the performance of the three models presented in Section 3.2 under the average cost criterion.

Example 3.4.1. Consider a system with two customer classes. The amount of resources available to this system is, $S = 10$ and the system incurs cost according to a quadratic utilization cost function $c(s) = \frac{s^2}{2}$. We compared the performance of the three models under 16 scenarios which differ based on arrival rates and holding costs. Input parameters are provided in Table 3.1. A numerical study comparing the performance of the three models under long-run average cost criterion is also shown in Table 3.1. For the systems modeled as MDPs, we truncate the state space so that each queue has a capacity of 50 and use the policy iteration algorithm to compute the approximate optimal policy.

Ex.	λ_1	λ_2	μ_1	μ_2	h_1	h_2	Cost full flex.	Cost lim flex.	Per diff	Cost Static	Per diff
1	2	2	1	1	2	1	13.33	13.54	1.56	30.87	131.62
2	2	2	1	1	5	1	14.81	15.37	3.81	41.19	178.17
3	2	2	1	1	10	1	16.84	18.00	6.86	54.47	223.44
4	2	2	1	1	15	1	18.57	20.39	9.78	58.16	213.21
5	0.5	3.5	1	1	2	1	12.86	13.23	2.85	38.87	202.16
6	0.5	3.5	1	1	5	1	13.15	14.01	6.52	44.37	237.39
7	0.5	3.5	1	1	10	1	13.57	15.01	10.62	52.26	285.06
8	0.5	3.5	1	1	15	1	13.95	15.87	13.78	54.15	288.31
9	3	1	1	1	2	1	13.88	14.68	5.73	37.73	171.76
10	3	1	1	1	5	1	16.60	17.59	6.01	52.71	217.58
11	3	1	1	1	10	1	20.09	21.49	6.99	58.58	191.66
12	3	1	1	1	15	1	22.97	24.86	8.21	61.57	168.05
13	1	1	1	1	2	1	4.94	4.95	0.02	16.35	230.78
14	1.5	1.5	1	1	2	1	8.68	8.71	0.33	23.11	166.19
15	2	2	1	1	2	1	13.33	13.54	1.56	30.87	131.62
16	3	3	1	1	2	1	25.42	26.9	5.85	49.38	94.29

Table 3.1: Comparison of average costs for the three systems in Example 3.4.1

The system with full flexibility should perform the best. However the fixed allocation system is desirable from the implementation standpoint. Our numerical experiment makes it clear that the Fixed allocation scheme is highly sub-optimal. This makes it clear that although the fixed allocation model is easier to implement in practice, substantial gains can be made by adopting a more flexible mechanism. Limited flexibility model performs well especially in low-med. load scenarios and when the queues are similar (ex 1, 13 and 14). However its performance starts to degrade as the load increases or when the holding cost at station 1 increases.



(a) Low Load (40%) and Increasing Holding Costs.

(b) Increasing Load (20% - 60%).

Figure 3.2: Breakdown of total costs for the first 8 scenarios discussed in the Example 3.4.1

A breakdown of the total costs for the three systems (Figure 3.2) shows that the limited flexibility system incurs lesser utilization cost than the fully flexible system. Although the fully flexible system incurs slightly more holding cost at station 2, it controls the holding cost at station 1 (the more expensive of the two) more tightly.

3.5 Heuristic Policies

When the number of stations in the models discussed in this paper become large, the computation of optimal control policies via standard iterative methods like value and policy iteration becomes prohibitive. In this section, we develop efficient heuristic methods that are used to develop policies for the system with full flexibility presented in Section 3.2. To test the performance of the dynamic programming solutions obtained in the previous sections, we compare the performance of the optimal control policy with the proposed heuristics and other widely used heuristics in the literature.

3.5.1 Performance Measures for a Priority Queue

In formulating the heuristic policies we make use of the long-run performance measures for a priority queueing system and we start this discussion by presenting some of these results. Consider a queueing system with m customer classes in which jobs of type k arrive according to a Poisson process with rate λ_k. The jobs are serviced by a single server and the service time for a class k customer is exponentially distributed with rate μ_k. Jobs are served according to a FCFS preemptive-resume priority service discipline. As per our convention, the classes are numbered such that the smaller the number higher is its priority. Define

$$\rho_i = \frac{\lambda_i}{\mu_i} \quad i = 1, 2, \ldots, m, \tag{3.35}$$

to be the occupation rate for class i i.e, the fraction of time the server spends serving class i customers in the long-run. The total load seen by the system is

$$\rho = \sum_{i=1}^{m} \rho_i.$$

Such a system has a unique stationary distribution if $\rho < 1$. Let Q_i denote the average number of class i customers present in the system in the long-run. The expression for Q_i can be written in terms of the occupation rates (see e.g, [25])

$$Q_i = \frac{\rho_i(1 + \sum_{n=1}^{i-1}(\frac{\mu_n}{\mu_i})\rho_n)}{(1 - \sum_{n=1}^{i-1}\rho_n)(1 - \sum_{n=1}^{i}\rho_n)} \quad i = 1, 2, \ldots, m. \tag{3.36}$$

3.5.2 Simple Heuristics

We next describe three simple heuristics based on the long-run performance measure for a priority queue.

1. *Static Rate Policy (SR):* The simplest heuristic for our system is one that selects a fixed number of servers s^f to operate at all times. The number of servers to operate are optimally selected to minimize the long-run total cost incurred by the system. The controller schedules the service at various stations according to the $c-\mu$ pre-emptive priority rule. The optimal capacity to be operated is selected by solving the following convex optimization problem:

$$\begin{aligned} \text{minimize} \quad & c(s) + \sum_{i=1}^{m} h_i Q_i^s(s) \\ \text{subject to} \quad & \sum_{i=1}^{m} \rho_i^s(s) < 1 \\ & 0 \le s \le S, \end{aligned} \tag{3.37}$$

   where $Q_i^s(s)$ are the number of customers of type i in a priority queueing system when the service rate at the i^{th} queue is $s\mu_i$ and $\rho_i^s(s)$ is the corresponding occupation rate. These are given (3.36) and (3.35) respectively (with μ_i replaced by $s\mu_i$).

2. *Gated - Static Rate Policy (GR):* The gated rate policy like the static rate policy selects a fixed number of servers, s^{gf}, to operate whenever there are jobs in the system. However, when the system is empty, the controller sets the service capacity to zero. As in the static policy case, the controller schedules the service at various stations according to the $c-\mu$ pre-emptive priority rule. The optimal capacity to operate when there are jobs to be

served is computed by solving the following convex optimization problem:

$$\begin{aligned} \text{minimize} \quad & \Big(\sum_{i=1}^{m} \rho_i^g(s)\Big) c(s) + \sum_{i=1}^{m} h_i Q_i^g(s) \\ \text{subject to} \quad & \sum_{i=1}^{m} \rho_i^g(s) < 1 \\ & 0 \le s \le S, \end{aligned} \tag{3.38}$$

where $Q_i^g(s)$ and $\rho_i^g(s)$ are given by (3.36) and (3.35) respectively (with μ_i replaced by $s\mu_i$).

3. *Gated - Station Dependent Policy (GS):* This policy like the gated-static rate policy sets the service rate to zero when there are not jobs in the system. However it selects a station dependent fixed number of servers s_i^{gs} when serving at station i. The controller schedules the service at stations based on the $c-\mu$ pre-emptive priority rule. The capacity selected when serving at station i is obtained by solving the following convex optimization problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{m} \rho_i^{gs}(s_i) c(s_i) + \sum_{i=1}^{m} h_i Q_i^{gs}(s) \\ \text{subject to} \quad & \sum_{i=1}^{m} \rho_i^{gs}(s_i) < 1 \\ & 0 \le s_i \le S, \quad i = 1, 2, \ldots, m, \end{aligned} \tag{3.39}$$

where $Q_i^{gs}(s)$ and $\rho_i^{gs}(s)$ are given by (3.36) and (3.35) respectively (with μ_i replaced by $s_i\mu_i$).

3.5.3 Single Queue (SQ) Heuristic

The main obstacle in computing the optimal service capacity for systems with multiple customer classes is that we need to keep track of the number of customers at each station. Since the state space grows exponentially in the number of stations, numerical computation of optimal controls becomes prohibitive for $m \geq 2$. The single queue based heuristic schedules the jobs according to the $c - \mu$ pre-emptive priority policy. It mitigates the problem of computing service capacity by approximating the system as a single class M/M/1 queue with an appropriate holding cost rate ($\bar{h}$), arrival rate ($\bar{\lambda}$) and service rate ($\bar{\mu}$). Based on the total number in the system, we compute the service capacity by solving a service rate control problem for this single queue. The state dependent service policy for such systems can be computed efficiently using the numerical procedure proposed by George and Harrison [23]. We still need to answer the question as to how to compute the input parameters ($\bar{h}$, $\bar{\lambda}, \bar{\mu}$) for the single queue heuristic. The natural choice for arrival rate is $\bar{\lambda} = \sum_{i=1}^{m} \lambda_i$. A proxy for the service rate is

$$\bar{\mu} = \frac{\sum_{i=1}^{m} \lambda_i \mu_i}{\sum_{i=1}^{m} \lambda_i}. \tag{3.40}$$

The question of how to select the holding cost rate is still a challenging one. For this we need to consider the contribution to holding cost from all m stations of the original system. One way to do this is by computing a weighted holding cost using the performance measures for the *Gated - Station Dependent Policy*. In particular, let Q_i^{gs} be the average number of class i jobs in the system under the *Gated - Station Dependent Policy* (this can be computed by solving (3.39)). Define

$$w_i = \frac{Q_i^{gs}}{\sum_{i=1}^{m} Q_i^{gs}} \tag{3.41}$$

as the fraction of class i jobs in the system in long-run under the *Gated - Station Dependent Policy*. Approximate the holding cost rate for the single queue heuristic using the weighted average of holding cost rates of the original system based on the weights w_i i.e,

$$\bar{h} = \sum_{i=1}^{m} w_i h_i. \tag{3.42}$$

Our numerical study, shows that this choice of parameters works extremely well.

3.5.4 Numerical Study

A numerical study comparing the performance of these heuristics is provided in Examples 3.5.1 3.5.2 and 3.5.3. In example 3.5.1 we show these results when the number of stations, $m = 2$. In this case we computed the performance of the optimal policy and of various heuristics by truncating the state space and using the policy iteration method. Example 3.5.2 deals with the case with 3 stations. Since the state space is quite large even for low queue truncation levels, we were able to compute the optimal policy only for low load cases. For moderate and high load scenarios we resort to a simulation study in Example 3.5.3 and compare the performance of the single queue heuristic with the gated (GR & GS) and static rate (SR) heuristics. In all these examples we generated a wide variety of test cases for low, medium and high loads. For each of these loads, we consider queues that differ based on their holding costs, arrival rates and service rates.

Example 3.5.1. Consider a system with two classes of jobs ($m = 2$). The total service capacity available to the system is, $S = 10$ and the cost of using s resources is $c(s) = \frac{s^2}{2}$. The arrival rates, service rates and holding costs for 20 different scenarios are provided in Table 3.2. Recall that the load at station i is denoted by $\Lambda_i = \frac{\lambda_i}{\mu_i}$ and the total load in the system by $\Lambda = \sum_{i=1}^{m} \Lambda_i$. We denote the system utilization by $\rho = \Lambda/S$. The queue lengths for each of these queues are restricted to 80. A comparison of the long-run average cost incurred by the optimal policy and various heuristics is shown in Table 3.2.

Ex.		$\frac{\Lambda_1}{\Lambda_2}$	μ_1	μ_2	h_1	h_2	Opt	SR	% Sub	GS	% Sub	GR	% Sub	SQ	%Sub
1	$\rho = 0.3$	0.3	1	1	2	1	8.45	11.24	33.07	8.95	5.97	9.01	6.54	8.46	0.12
2		0.5	1	1	2	1	8.68	11.56	33.17	9.16	5.45	9.25	6.53	8.71	0.25
3		0.7	1	1	2	1	9.02	12.09	34.14	9.48	5.16	9.62	6.64	9.05	0.41
4		0.3	2	1	5	1	9.01	11.78	30.9	9.45	4.98	9.57	6.36	9.11	1.06
5		0.5	2	1	5	1	9.79	12.92	31.94	10.21	4.19	10.48	7.06	9.99	2.01
6		0.7	2	1	5	1	10.87	14.91	37.09	11.46	5.41	11.73	7.91	11.15	2.56
7	$\rho = 0.5$	0.3	1	1	2	1	18.58	23.3	25.38	19.81	6.63	19.88	6.96	18.59	0.04
8		0.5	1	1	2	1	18.9	23.69	25.34	20.08	6.23	20.21	6.92	18.92	0.11
9		0.7	1	1	2	1	19.39	24.39	25.81	20.59	6.18	20.75	7.02	19.43	0.24
10		0.3	2	1	5	1	19.18	23.66	23.33	20.32	5.9	20.41	6.41	19.25	0.36
11		0.5	2	1	5	1	20.24	24.88	22.93	21.3	5.23	21.56	6.53	20.4	0.82
12		0.7	2	1	5	1	21.91	27.55	25.77	23.26	6.22	23.58	7.65	22.25	1.61
13	$\rho = 0.7$	0.3	1	1	2	1	32.55	39.33	20.84	34.66	6.49	34.73	6.71	32.55	0.02
14		0.5	1	1	2	1	32.92	39.76	20.77	34.99	6.28	35.11	6.65	32.94	0.06
15		0.7	1	1	2	1	33.54	40.57	20.96	35.64	6.27	35.78	6.67	33.58	0.14
16		0.3	2	1	5	1	33.09	39.44	19.2	35.07	5.98	35.14	6.21	33.17	0.26
17		0.5	2	1	5	1	34.31	40.61	18.36	36.17	5.44	36.36	5.99	34.41	0.31
18		0.7	2	1	5	1	36.53	43.61	19.4	38.65	5.81	38.9	6.52	36.83	0.84

Table 3.2: Average Cost Rates and Percentage Difference between the Optimal and Heuristic Policies for Example 3.5.1

Table 3.2 shows that the single queue heuristic (SQ) performs near optimally in all test cases. This heuristic is within 1% of the optimal cost in most cases and around 3% sub-optimal in the worst case. Both gated heuristics (GR & GS) also perform quite well resulting in policies that are within $(5 - 8\%)$ of the optimal cost. The performance of static rate policies range from 19% to 30% sub-optimal, this shows that there is substantial benefit in investing in more responsive mechanisms.

Example 3.5.2. In this example, we consider a system with three customer classes ($m = 3$). The total service capacity available to the system is, $S = 10$ and the cost of using s resources is $c(s) = s^3$. The queue lengths for each of these queues were restricted to 30 and policy iteration algorithm was used to compute the optimal policy. Due to the low truncation levels for the queues, we could reliably find the optimal policy only in low load (10 – 30%) scenarios. The arrival rates, service rates and holding costs for 9 such scenarios are given in Table 3.3. A comparison of the long-run average cost incurred by the optimal policy and various heuristics is shown in Table 3.4.

Ex		λ_1	λ_2	λ_3	μ_1	μ_2	μ_3	h_1	h_2	h_3
1		0.15	0.15	0.7	1	1	1	3	2	1
2	$\rho = 0.1$	0.25	0.25	0.5	1	1	1	3	2	1
3		0.35	0.35	0.3	1	1	1	3	2	1
4		0.7	0.7	0.6	1	1	1	3	2	1
5	$\rho = 0.2$	0.3	0.3	1.4	1	1	1	3	2	1
6		0.5	0.5	1	1	1	1	3	2	1
7		1.05	1.05	0.9	1	1	1	3	2	1
8	$\rho = 0.3$	0.75	0.75	1.5	1	1	1	3	2	1
9		0.45	0.45	2.1	1	1	1	3	2	1

Table 3.3: Input parameters for Example 3.5.2

Ex	Opt	SR	% Sub	GR	% Sub	GS	% Sub	SQ	% Sub
1	2.02	3.16	56.33	2.12	4.66	2.08	2.96	2.04	1.01
2	2.16	3.41	57.86	2.27	4.87	2.21	2.43	2.19	1.4
3	2.32	3.73	60.45	2.43	4.71	2.37	2.14	2.35	1.41
4	5.39	7.67	42.24	5.72	6.03	5.59	3.69	5.44	0.95
5	4.85	6.8	40.2	5.13	5.88	5.08	4.79	4.87	0.45
6	5.09	7.16	40.66	5.4	5.97	5.3	4.12	5.13	0.77
7	9.29	12.46	34.09	9.91	6.66	9.74	4.81	9.35	0.69
8	8.87	11.8	33.06	9.45	6.55	9.33	5.15	8.91	0.46
9	8.55	11.37	32.95	9.11	6.54	9.05	5.79	8.57	0.24

Table 3.4: Average Cost Rates and Percentage Difference between the Optimal and Heuristic Policies for Example 3.5.2

We again see that the single queue (SQ) heuristic performs near optimally (with-in 2% of optimal) in all cases. Both gated heuristic policies (GR & GS) perform well and are within 7% of the optimal cost. Note that in Table 3.3, the loads are increasing from 10% to 30% as we move from case 1 to 9. We see that performance of gated policies degrades as the load increases. The opposite is true for the static rate heuristic with its performance improving as the load increases.

Example 3.5.3. In this example we compare the performance of the single queue heuristic with other heuristics using an in-depth simulation study. We again consider the 3 station case presented in Example 3.5.2 but now for moderate loads (up to 50 %). The input parameters considered in this study are presented in Table 3.5. For each of these cases, a simulation run consisted of 50,000 jobs completions from the system and each run was repeated 24 times. The average long-run cost over these runs and the corresponding 95% confidence intervals for various cases are shown in Table 3.6.

Ex.	λ_1	λ_2	λ_3	μ_1	μ_2	μ_3	h_1	h_2	h_3
1	1.05	1.05	0.9	1	1	1	3	2	1
2	0.75	0.75	1.5	1	1	1	3	2	1
3	1.75	1.75	1.5	1	1	1	3	2	1
4	1.25	1.25	2.5	1	1	1	3	2	1
5	0.75	0.75	3.5	1	1	1	3	2	1
6	1.05	1.05	0.9	1	1	2	4	3	1
7	0.75	0.75	1.5	1	1	2	4	3	1
8	0.45	0.45	2.1	1	1	2	4	3	1
9	1.75	1.75	1.5	1	1	2	4	3	1
10	1.25	1.25	2.5	1	1	2	4	3	1
11	0.75	0.75	3.5	1	1	2	4	3	1
12	2.45	2.45	2.1	1	1	2	4	3	1
13	1.75	1.75	3.5	1	1	2	4	3	1
14	1.05	1.05	4.9	1	1	2	4	3	1
15	1.05	1.05	0.9	1	2	2	7	2	1
16	0.75	0.75	1.5	1	2	2	7	2	1
17	0.45	0.45	2.1	1	2	2	7	2	1
18	1.75	1.75	1.5	1	2	2	7	2	1
19	1.25	1.25	2.5	1	2	2	7	2	1
20	0.75	0.75	3.5	1	2	2	7	2	1
21	2.45	2.45	2.1	1	2	2	7	2	1
22	1.75	1.75	3.5	1	2	2	7	2	1
23	1.05	1.05	4.9	1	2	2	7	2	1

Table 3.5: Input parameters for Example 3.5.2

Ex.	SQ	GS	% Sub	GR	% Sub	SR	% Sub
1	40.67 ± 0.22	43.73 ± 0.29	7.52	44.45 ± 0.26	9.29	48.24 ± 0.34	18.61
2	39.57 ± 0.25	43.18 ± 0.33	9.12	43.55 ± 0.42	10.06	46.55 ± 0.32	17.64
3	148.41 ± 1.01	158.56 ± 1.07	6.84	158.44 ± 0.89	6.76	166.94 ± 1.56	12.49
4	146.98 ± 0.89	156.81 ± 0.96	6.69	158.41 ± 1.12	7.78	166.02 ± 1.24	12.95
5	146.75 ± 0.88	156.35 ± 1.02	6.54	157.03 ± 1.11	7.01	165.06 ± 1.22	12.48
6	32.51 ± 0.19	34.81 ± 0.21	7.07	35.46 ± 0.17	9.07	40.25 ± 0.24	23.81
7	23.76 ± 0.13	25.36 ± 0.13	6.73	25.91 ± 0.16	9.05	29.46 ± 0.21	23.99
8	16.62 ± 0.1	17.74 ± 0.12	6.74	18.13 ± 0.1	9.09	20.99 ± 0.13	26.29
9	105.19 ± 0.63	113.88 ± 0.79	8.26	114.51 ± 0.71	8.86	123.77 ± 0.71	17.66
10	74.7 ± 0.48	80.73 ± 0.63	8.07	81.71 ± 0.5	9.38	89.12 ± 0.69	19.3
11	51.25 ± 0.35	55.3 ± 0.38	7.9	55.97 ± 0.45	9.21	60.73 ± 0.41	18.5
12	253.82 ± 2.07	268.97 ± 2.08	5.97	269.39 ± 1.39	6.13	283.09 ± 1.2	11.53
13	178.12 ± 1.08	189.47 ± 1.13	6.37	190.62 ± 1.18	7.02	199.91 ± 1.19	12.23
14	119.26 ± 0.84	127.75 ± 0.84	7.12	127.94 ± 0.71	7.28	137.27 ± 1.35	15.1
15	21.7 ± 0.12	22.74 ± 0.14	4.79	23.4 ± 0.13	7.83	27.1 ± 0.16	24.88
16	16.98 ± 0.09	17.8 ± 0.1	4.83	18.6 ± 0.09	9.54	21.3 ± 0.14	25.44
17	13.15 ± 0.07	13.94 ± 0.07	6.01	14.27 ± 0.08	8.52	16.6 ± 0.09	26.24
18	61.28 ± 0.27	65.72 ± 0.48	7.25	66.81 ± 0.44	9.02	72.78 ± 0.43	18.77
19	48.52 ± 0.25	52.54 ± 0.42	8.29	52.94 ± 0.29	9.11	58.03 ± 0.55	19.6
20	37.88 ± 0.23	41.22 ± 0.25	8.82	41.7 ± 0.25	10.08	45.37 ± 0.27	19.77
21	137.71 ± 0.98	146.89 ± 1	6.67	147.4 ± 0.94	7.04	157.46 ± 1.21	14.34
22	109.78 ± 0.75	118.02 ± 0.76	7.51	118.31 ± 0.76	7.77	126.27 ± 1.02	15.02
23	85.76 ± 0.56	93.2 ± 0.67	8.68	92.81 ± 0.57	8.22	99.43 ± 0.9	15.94

Table 3.6: Estimated mean and 95% confidence interval for the long-run average costs in Example 3.5.3

Table 3.6 shows that the single queue (SQ) heuristic outperforms the gated and static rate heuristics in all cases. In particular, compared to the SQ heuristic, the performance loss of Gated - Station Dependent policy (GS) is between 4 – 10%, that of Gated - Static Rate policy (GS) is between 6 – 10% while that of static rate policy is between 12 – 27%.

3.6 Conclusions

In this Chapter we investigate the dynamic resource allocation problem for the model with multiple job types. Assuming that the goal is to minimize a combination of effort cost and holding cost, we study this problem under both the discounted and average cost optimality criterion. We develop two dynamic allocation models based on varying degree of flexibility. For both these models, we characterize the structure of the optimal policy. The optimal scheduling policy in both cases is a $c - \mu$ priority scheme. We also show that the resource utilization levels are non-decreasing in the number of jobs of each type. When the number of stations are large, finding optimal policies for the full model may become difficult. For such scenarios we develop several heuristics and compare their performance using numerical studies. In particular, we find that the Single Queue heuristic presented in Section 3.5 performs very well for a large class of examples.

CHAPTER 4

DYNAMIC CAPACITY MANAGEMENT FOR A MULTI-CLASS QUEUE WITH SET-UP COSTS

4.1 Introduction

A drawback of the systems discussed in Chapter 3 is that the servers can be reallocated to various stations instantaneously. In practice, most systems incur some set-up time or set-up cost when resources are switched from one station to another. In this section, we address the scenario when the system incurs a cost penalty when resources switch between stations. We continue to assume that the both switches and changes in resource capacity are instantaneous. Although we are ignoring switch-over times, a model with just switching costs can still be beneficial from the perspective of finding policies that are easy to implement. Implementing policies for the fully flexible system presented in Chapter 3 can be difficult due to frequent switching of servers between stations, incorporating appropriate switching costs can mitigate this problem.

Queueing models with set-up costs or times are challenging to analyze and the presence of utilization cost in our model makes the task even more difficult. Most of the research in this area has focussed on models with only one server and without utilization costs. Hofri et al. [30] and Liu et al. [37], studied a model with two homogeneous parallel queues in the presence of both set-up costs and set-up times. They characterized the optimal policy as being exhaustive at each station and conjectured that switching from an empty queue follows a threshold policy. Duenyas et al. [18] and Koole et al. [34] studied the problem with two customer classes in the presence of set-up costs. They provide a partial characterization of the optimal policy as being exhaustive at the queue with the highest $c-\mu$ index. Duenyas et al. [19] provide a similar characterization for a two class queueing model but with general service time distribution. Reiman et al. [47] study a similar queueing system under the heavy traffic approximation. In this case they completely characterize the nature of the optimal policy. For the case of multiple servers, Sledgers et al. [51] consider a multi-class queueing model with Markov modulated arrivals and set-ups. They develop an MDP based model and provide several heuristics for this model. To the best of our knowledge, none of these works consider the model with utilization costs.

4.2 Preliminaries and Model Description

Consider the model presented in Chapter 3, Section 3.2 for the system with full flexibility. The system has access to S identical resources each of which can be allocated to one of the m stations. A central controller constantly monitors the number of pending jobs of each type present in the system and dynamically decides the number of resources to be set-up at station i and how many from those set-up to use at each station. The servers can be switched between stations and be turned on/off at any time; and such changes are instantaneous. Note that since our model incurs a utilization cost, it may be optimal to idle some of the capacity set-up for a particular station. We are considering that resources can be chosen from a discrete capacity set $\{0, 1, \ldots, S\}$. We next formulate this problem as a Markov decision process.

Let x_i denote the total number of jobs at station i (including the job being processed). We denote the number of servers currently set-up at station i by u_i.

We consider direction dependent switching costs in this model. This addition requires us to augment state space by incorporating the number of servers set-up at each station. Let $\mathbb{X} = \mathbb{Z}_+^m$ and $\mathbb{U} = \{(u_1, u_2, \ldots, u_m) | u_i \in \{0, 1, \ldots, S\}, \sum_{i=1}^m u_i = S\}$. The state space for this system is $\mathbb{X}^S = \mathbb{X} \times \mathbb{U}$. The action set is comprised of two decisions; the number of servers to switch from station i to j denoted by ℓ_{ij} and how many out of the available servers at station i to utilize for service denoted by s_i. Thus the action space in state (x, u) is

$$\mathbb{A}(x,u) = \{(\ell, s) : \sum_{j \neq i} \ell_{ij} \leq u_i, \ell_{ij} \in \{0, 1, \ldots, u_i\}; s_i \in \{0, 1, \ldots, u_i + \sum_{j \neq i} \ell_{ji} - \sum_{j \neq i} \ell_{ij}\}, i = 1, 2, \ldots, m\}. \tag{4.1}$$

Note that if ℓ_{ij} servers are switched between stations i and j in state (x, u), the state immediately changes to (x, u') where

$$u_i' = u_i + \sum_{j \neq i} \ell_{ji} - \sum_{j \neq i} \ell_{ij}, \quad i = 1, 2, \ldots, m. \tag{4.2}$$

Furthermore, the system incurs a direction dependent, non-negative switching cost R_{ij} per server switched from the station i to j. We assume that the switching costs satisfy the triangle inequality i.e, for any pair (i, j), $R_{i,j} \geq R_{i,k} + R_{k,j}, k = 1, 2, \ldots, m$. This means that only direct switches between stations are optimal and the action space considered in (4.1) is sufficient. Furthermore, since the set-up costs are non-negative, only uni-directional switches between any pair of stations are optimal i.e, for any pair of stations i, j, with $\ell_{ij} + \ell_{ji} > 0$, either $\ell_{ij} > 0$ or $\ell_{ji} > 0$. The decision epochs are the job arrival, service completion and dummy transition times (due to uniformization described later). When the system is in state $(x, u) \in \mathbb{X}^S$ and action $(\ell, s) \in \mathbb{A}(x, u)$ is chosen, the system evolves with the following transition rates

$$q(((x,u),(\ell,s)),(y,u')) = \begin{cases} \lambda_i & \text{if } y = x + e_i, i = 1,\dots,m \\ s_i\mu_i & \text{if } x_i > 0, y = (x - e_i)^+, i = 1,\dots,m \\ 0 & \text{otherwise,} \end{cases} \tag{4.3}$$

where u' is given by (4.2). Let $K(\ell, s)$ denote the total switching cost incurred under action (ℓ, s). Then

$$K(\ell, s) = \sum_{i=1}^{m}\sum_{j=1}^{m} R_{i,j}\ell_{i,j}. \tag{4.4}$$

Let $a(\ell, s) = \sum_{i=1}^{m} s_i$, denote the total number of active servers at all stations under action (ℓ, s). The system incurs the following cost rate in state (x, u) under action (ℓ, s)

$$C((x,u),(\ell,s)) = c(a(\ell,s)) + \sum_{i=1}^{m} h_i x_i. \tag{4.5}$$

Let Π denote the set of non-anticipating policies, and suppose for any $\pi \in \Pi$ that $\{(x^\pi(t), u^\pi(t), (\ell^\pi(t), s^\pi(t))); t \geq 0\}$ denote the stochastic process representing the evolution of state and decisions under policy π. Also, let $L^\pi_{ij}(t)$ count the total number servers that have switched from stations i and j until time t. Given the initial state (x, u) and discount factor $\alpha > 0$, the $\alpha-$**discounted finite horizon expected cost**, under policy π is given by

$$v^\pi_{T,\alpha}(x,u) := \mathbb{E}^\pi_{(x,u)}\left[\int_0^T e^{-\alpha t} C((x^\pi(t), u^\pi(t)), (\ell^\pi(t), s^\pi(t)))dt + \int_0^T e^{-\alpha t}\sum_{i=1}^{m}\sum_{j=1}^{m} R_{ij} dL^\pi_{ij}(t)\right], \tag{4.6}$$

where $\mathbb{E}^\pi_{x,u}[\cdot]$ is the expectation with respect to the law of the stochastic process induced by the policy π when the starting state is (x, u). The **total discounted expected cost** under a policy π given that the initial state of the system is (x, u), is $v^\pi_\alpha(x, u) := \lim_{T\to\infty} v^\pi_{T,\alpha}(x, u)$ and the **long-run average cost** is

$g^\pi(x,u) := \limsup_{T\to\infty} \frac{v^\pi_{T,0}(x,u)}{t}$. In each case, the optimal costs starting from state (x,u) are

$$f^*(x,u) := \inf_{\pi\in\Pi} f^\pi(x,u), \tag{4.7}$$

for $f = v_\alpha$ or g and the policy, π^*, that attains the infimum for all $(x,u) \in \mathbb{X}^S$, is optimal in the respective case. Our objective is to find an allocation policy π that minimizes the total expected discounted cost or the expected average cost rate incurred by the system over an infinite planning horizon.

Since the holding rates under any admissible policy are bounded, we can apply *uniformization* in the spirit of Lippman [36] and consider the discrete time equivalent of the continuous time Markov chain described above. The uniformization rate is chosen to be $\gamma := \sum_{i=1}^m \lambda_i + S\bar{\mu}$. Without loss of generality let $\gamma = 1$. Note that uniformization induces a discount factor $\beta = \frac{\gamma}{\alpha+\gamma}$ for the discrete time chain.

Define $v_{\alpha,k}(x,u)$ be the minimum total α-discounted expected cost that can be obtained during the last k transitions when starting from state (x,u) and let $v_{\alpha,0}(x,u) = 0$ for all $(x,u) \in \mathbb{X}^S$. Using standard arguments of Markov decision theory [11], the discrete-time finite horizon optimality equations (FHOE) for the system can be written

$$v_{\alpha,k+1}(x,u) = \mathcal{T}^S_\alpha v_{\alpha,k}(x,u) \tag{4.8}$$

$$= \min_{(\ell,s)\in\mathbb{A}(x,u)} \Big\{ K(\ell,s) + \frac{C((x,u),(\ell,s))}{\alpha+1} + \beta\Big(\sum_{i=1}^m \lambda_i v_{\alpha,k}(x+e_i,u') + \sum_{i=1}^m s_i\mu_i v_{\alpha,k}((x-e_i)^+,u')$$

$$+ (1 - \sum_{i=1}^m (\lambda_i + \mathbb{1}_{\{x_i>0\}} s_i\mu_i)) v_{\alpha,k}(x,u')\Big)\Big\} \tag{4.9}$$

where the updated partial state, u' is given by (4.2) and e_i denotes the vector of all zeros except one in the i^{th} coordinate. The set-up cost $K(\ell, s)$ and cost rate $C((x, u), (\ell, s))$ in (4.9) are given in (4.4) and (4.5) respectively. $\mathcal{T}_\alpha^S$ in (4.9) is the dynamic programming mapping defined on real valued functions on $\mathbb{X}^S$. Note that we may apply Theorem 6.10.4 of [46] to get $v_\alpha^k \uparrow v_\alpha$. In particular let $r(x, u) := \sum_{i=1}^m h_i x_i + c(S\bar{\mu}) + S \max_{i,j} R_{ij}$ and $v_r := \sup_{(x,u)\in\mathbb{X}^S} r(x, u)^{-1}|v(x, u)|$ be the weighted supremum norm for real-valued functions v on $\mathbb{X}^S$. Then the cost rate function $C(x, (a, s))$ and transitions rates $q((x, (a, s)), y)$ defined in (4.3) satisfy Assumptions 6.10.1 and 6.10.2 of [46] (take $\mu = 1$ for 6.10.1 and $L = \sum_{i=1}^m h_i\lambda_i$ for 6.10.5 (a)). Moreover, v_α satisfies the *discounted cost optimality equations* (DCOE) stated below

$$v_\alpha(x, u) = \mathcal{T}_\alpha^S v_\alpha(x, u). \tag{4.10}$$

The average cost optimality equations for this system can be written as

$$g + w(x, u) = \mathcal{T}_0^S w(x, u). \tag{4.11}$$

When a solution (w, g) to the ACOE (4.11) exists, it is such that g is the optimal average cost (independent of the initial state) and the relative value function, $w(x, u)$, can be interpreted as the minimum expected cost incurred until the next entry into an arbitrary reference state $(x_{\text{ref}}, u_{\text{ref}}) \in \mathbb{X}^S$.

4.2.1 Numerical Solution

The set-up cost problem presents significant computational challenges. For example even a small sized problem with 5 servers, 3 stations and truncation level of 20 has 194481 states. Therefore any simplification in the model can be beneficial for speedy computations. In this section we illustrate several interesting and perhaps surprising features of the optimal policy using two examples. The simple examples presented next will help us in determining structures that can be exploited to build simpler models and heuristics with good performance. To make the intuition clear and keep the exposition simple, both problems deal with a system with two stations and two servers. In all the examples discussed here, we used policy iteration for solving the average cost problem defined in (4.11) and truncated the queue lengths to 40.

Example 4.2.1. Consider a system with two stations ($m = 2$). The total service capacity available to the system is, $S = 2$ and the cost of using s resources is $c(s) = \frac{s^2}{2}$. The arrival rates are $\lambda_1 = \lambda_2 = 0.1$ and the service rates are $\mu_1 = \mu_2 = 0.3$. The system is operating under low load condition with total load $(\sum_{i=1}^{m} \frac{\lambda_i}{S\mu_i}) = \frac{1}{3}$. The holding costs are $h_1 = 2$ and $h_2 = 1$. We first consider a fairly high switching cost with $R_{12} = R_{21} = 20$ per server switch. Note that station 1 is the more expensive station. The optimal average cost policy for this example is shown in Figure 4.1 for various queue lengths and server set-up configurations. The symbol corresponding to station state (i, j) denotes if server(s) should switch or stay at stations. In particular, (-) symbol denotes that server(s) should be switched from station 1 to station 2, (+) symbol denotes that server(s) should be switched from station 2 to station 1 and (o) symbol denotes that the servers should stay put at current station. The numbers below each symbol denotes the number of servers active at each station. For example, the optimal policy when $i = 2$, $j = 8$ and both servers are set-up at station 1 can be found by referring to Figure 4.1(a) which shows that the optimal action is to switch a server to station 2 and use one server each at these stations.

It is clear from Figure 4.1 that the general structure of the optimal policy is quite complex. However we can still make several important observations about the scheduling policy:

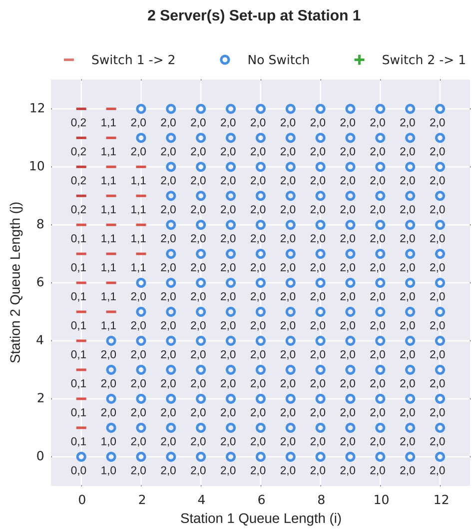
Observation 1 - Exhaustive scheduling at the most expensive queue is not optimal: If there are pending jobs at the most expensive station, station 1 in this case, it may be reasonable to argue that any server(s) set-up at this station should not be switched to the less expensive station 2. This is certainly true for the case without set-up costs (Theorem 3.3.10). Even for the case with set-up cost or set-up time for the problem with single server and no utilization cost, several authors have shown the optimality of the exhaustive policy (see for example [30],[18] and [19]). Figure 4.1 shows that this is not optimal for the current model. When two servers are set-up at station 2 and $i = 1$, $j = 6$, Figure 4.1(a) shows that in this case one of the servers set-up at station 1 should be switched to station 2.

Observation 2 - Policy for switching servers follows a switching curve but it is not monotone: When considering switching a server from station 1 to station 2, one might expect that if a server switch from 1 to 2 is optimal when $x = (i, j)$, same should hold in $x' = (i, j + 1)$ since the situation at station 2 is more severe. This structure is known as a monotone switching curve. However Figure 4.1(a) shows that it is optimal to switch a server from 1 to 2 when $i = 2$ and $j = 10$ while such a switch is not optimal when $i = 2$ and $j = 11$. Note that this is not a boundary effect due to truncation. In Figure 4.1 we show queue lengths only up to 12, our true truncation level is 40. This behaviour is also consistently seen in several other experiments we performed. Similar behavior turns out to be true for the policy that governs server switches from the low cost to the high cost station (e.g, see Figure 4.2(c)).

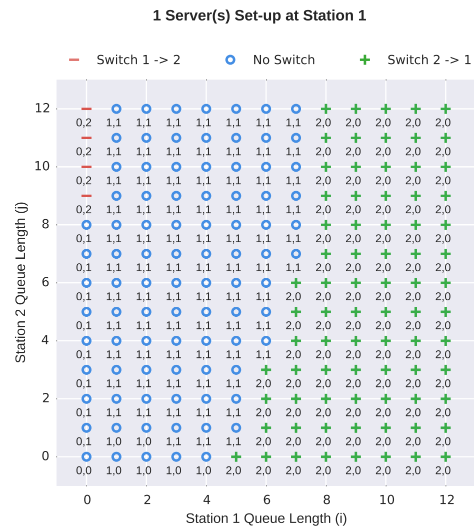
Observation 3 - Server collaboration is not optimal: In the model without set-up cost, we saw that servers always worked in a collaborative manner (Theorem 3.3.10). Not surprisingly this result no longer holds in the model with set-up costs. Figure 4.1 shows that in large number of states, it is optimal to stay in a split configuration (e.g, $i = 4$ and $j = 4$ in Figure 4.1(b)) or switch from a collaborative configuration to a split one (e.g, $i = 2$ and $j = 8$ in Figure 4.1(a)).

From the perspective of simplicity and computational tractability, policies in which the servers switch in a collaborative manner are desirable. Since we would only need to keep track of the station at which the servers are currently set-up in this case, we can achieve significant reduction in the state space. Unfortunately, Observation 3 in Example 4.2.1 shows that this is not optimal in general, especially when the set-up costs are high. A natural question is are there cases in which optimal policies switch in a collaborative fashion most of the time. Our next example shows that in the case when the total set-up costs are low, collaborative switching policies may not be far from optimal. This is perhaps not too surprising since we already know that in the model without set-up costs, collaborative switching is optimal (Theorem 3.3.10).

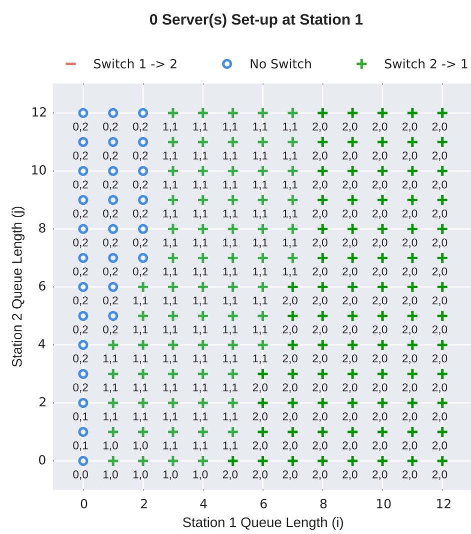
Example 4.2.2. We consider the same set-up as in Example 4.2.1. The only difference is that the switching costs are $R_{12} = R_{21} = 1$ per server switch. The optimal average cost policy for this example is shown in Figure 4.2 for various queue lengths and server set-up configurations. It is striking that in this example, most of the states with label $(1, 1)$ in Figure 4.1 have either become $(0, 2)$ or $(2, 0)$. This shows a strong preference for the servers to stay together and switch in a collaborative fashion under the optimal policy in this case.



(a) 2 servers set-up at station 1



(b) 1 server is set-up at station 1

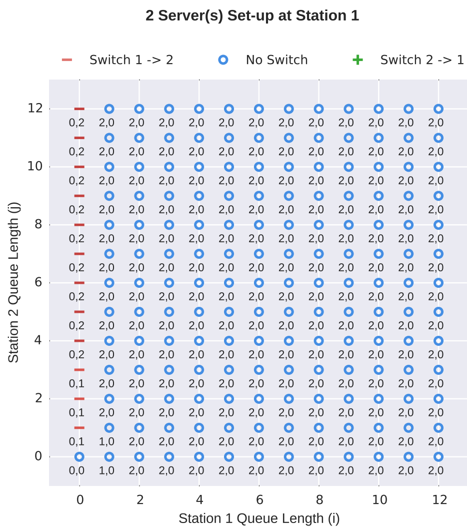


(c) no servers are set-up at station 1

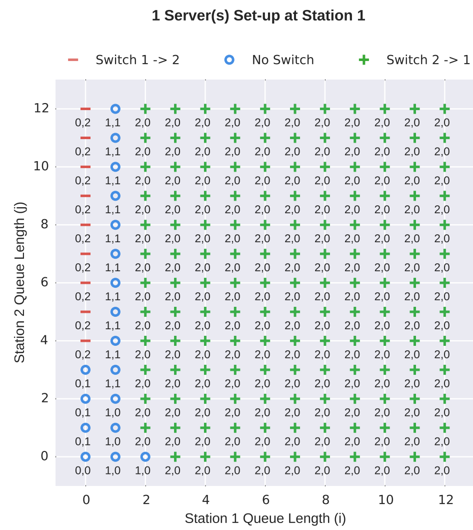
Figure 4.1: Optimal policy for medium set-up cost case discussed in example 4.2.1. Symbols represent if servers are switching away (-), towards (+) or staying (o) at station 1. Numbers below each symbol denotes the number of active servers at each station.

4.3 Model with Consolidated Switching

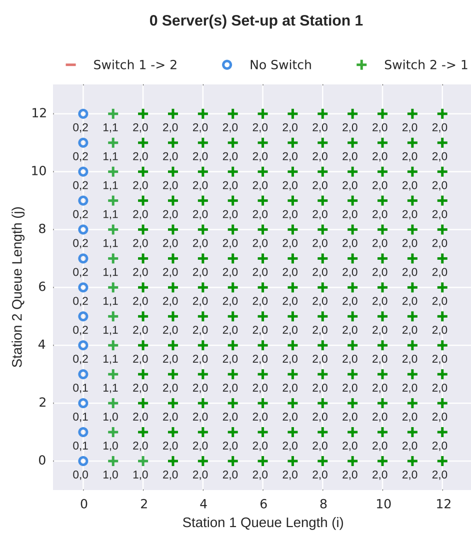
Example 4.2.2 displays that when set-up costs are not too high, collaborative switching policies may perform well. Models with consolidated switching are



(a) 2 servers are set-up at station 1



(b) 1 server is set-up at station 1



(c) no servers are set-up at station 1

Figure 4.2: Optimal policy for medium set-up cost case discussed in example 4.2.2. Symbols represent if servers are switching away (-), towards (+) or staying (o) at station 1. Numbers below each symbol denotes the number of active servers at each station.

desirable since these allow us to reduce our state space and compute optimal policies efficiently. In this section, we analyze the model when the servers can all be set-up for servicing jobs at only one station at any given time. Note that due to the utilization cost, not all of the servers set-up at a station may be active.

Apart from the restriction of collaborative switch, the set-up described in the Section 4.2 continues to hold.

The state of this system is given by the position of the server at any decision epoch and the number of customers present at each station, x. Let u denote the station for which servers are currently set-up. Denote the state of the system by (x,u) and the corresponding state space by $\mathbb{X}^{CS} = \mathbb{X} \times \{0,1,\ldots,m\}$. The action comprises of two decision; the station to work at, denoted by ℓ and how many servers to make active denoted by s. Denote this action pair by (ℓ, s) and the corresponding action space by $\mathbb{A} = \{0,1,2,\ldots,m\} \times \{0,1,2,\ldots,S\}$. The system incurs a direction dependent switching cost R_{ij} per server switched from the station i to j. Let $K((x,u),(\ell,s))$ denote the total switching cost incurred in state (x,u) under action (ℓ,s). Then since all servers switch together,

$$K((x,u),(\ell,s)) = SR_{u,\ell}. \tag{4.12}$$

The system incurs following cost rate in state (x,u) under action (ℓ,s)

$$C((x,u),(\ell,s)) = c(s) + \sum_{i=1}^{m} h_i x_i. \tag{4.13}$$

Suppose for any policy $\pi \in \Pi$ that $\{(x^\pi(t), u^\pi(t), (\ell^\pi(t), s^\pi(t))); t \geq 0\}$ denote the stochastic process representing the evolution of state and decisions under policy π. Also, let $L^\pi_{ij}(t)$ count the total number of switches made between stations i and j until time t. Given the initial state (x,u) and discount factor $\alpha > 0$, the $\alpha-$**discounted finite horizon expected cost**, under policy π is given by

$$v^\pi_{T,\alpha}(x,u) := \mathbb{E}^\pi_{(x,u)}\left[\int_0^T e^{-\alpha t} C((x^\pi(t), u^\pi(t)), (\ell^\pi(t), s^\pi(t)))dt + \int_0^T e^{-\alpha t} \sum_{i=1}^{m}\sum_{j=1}^{m} R_{ij} dL^\pi_{ij}(t)\right], \tag{4.14}$$

where $\mathbb{E}_{(x,u)}[\cdot] := \mathbb{E}[\cdot|x(0) = (x,u)]$. Define $v_{\alpha,k}(x,u)$ be the minimum total α-discounted expected cost that can be obtained during the last k transitions when starting from state (x,u) and let $v_{\alpha,0}(x,u) = 0$ for all $(x,u) \in \mathbb{X}$. Using standard arguments of Markov decision theory [11], the discrete-time finite horizon optimality equations (FHOE) for the system can be written

$$v_{\alpha,k+1}(x,u) = \mathcal{T}_{\alpha}^{CS} v_{\alpha,k}(x,u) \tag{4.15}$$

$$= \min_{(\ell,s)\in\mathbb{A}} \Big\{ S R_{u,\ell} + \frac{c(s) + \sum_{i=1}^{m} h_i x_i}{\alpha+1} + \beta\Big(\sum_{i=1}^{m} \lambda_i v_{\alpha,k}(x+e_i,\ell) + s_\ell \mu_\ell v_{\alpha,k}((x-e_\ell)^+,\ell)$$
$$+ (1 - \sum_{i=1}^{m} \lambda_i - \mathbb{1}_{\{x_\ell>0\}} s_\ell \mu_\ell) v_{\alpha,k}(x,\ell)\Big)\Big\} \tag{4.16}$$

where $\mathcal{T}_{\alpha}^{CS}$ is the dynamic programming mapping defined on real valued functions on the state space. As before, we may apply Theorem 6.10.4 of [46] to get $v_{\alpha}^{k} \uparrow v_{\alpha}$. Moreover, v_α satisfies the *discounted cost optimality equations* (DCOE) stated below

$$v_\alpha(x,u) = \mathcal{T}_{\alpha}^{CS} v_\alpha(x,u). \tag{4.17}$$

The average cost optimality equations for this system can be written as

$$g + w(x,u) = \mathcal{T}_{0}^{CS} w(x,u). \tag{4.18}$$

4.4 Structure of Optimal Policy for the Consolidated Switching Model

In this section, we partially characterize the optimal policy for the model with consolidated switching. To make the intuition behind the proof clear and keep the exposition simple, we present this result for the two station case ($m = 2$). The m station case is a simple extension of this result.

Theorem 4.4.1. *Let $h_1\mu_1 \geq h_2\mu_2$, then under the infinite horizon discounted cost criteria or the average cost criteria, there exists an optimal policy that serves at station 1 in an exhaustive manner.*

We prove this result via sample path methods and an interchange argument in a manner similar to [43] and [17]. Let $\{(x^\pi(t), u^\pi(t), (\ell^\pi(t), s^\pi(t))); t \geq 0\}$ denote the stochastic process representing the evolution of state and decisions under policy π. Recall that $a^\pi(t)$ denotes the total number of servers being used at time t and $L^\pi_{ij}(t)$ counts the total number of switches made between stations i and j until time t. For a policy π, define for any time t

$$\theta_\pi(t) := \int_0^t \Big[h_1\mu_1 a^\pi(y)\delta_1^\pi(y) + h_2\mu_2 a^\pi(y)\delta_2^\pi(y)\Big]dy, \tag{4.19}$$

where $\delta_i^\pi(y) := \mathbb{1}_{\{x_i^\pi(y^-)>0, d^\pi(y)=i\}}$ and $d^\pi(y) = i$ if all resources are allocated at station i at time y. By serving jobs, a policy π stands to save $\theta_\pi(t)$ amount of holding cost until time t.

The following lemma uses (4.19) to re-express the finite horizon discounted expected cost in terms of $\theta_\pi(t)$ that is more conducive to sample-path based analysis. Let $\{A_i(t); t \geq 0\}$ denote the cumulative arrival processes at station i.

Lemma 4.4.2. *The finite horizon discounted expected cost under policy $\pi \in \Pi$ is*

$$\begin{aligned} v^\pi_{t,\alpha} = {} & \frac{1 - e^{-\alpha t}}{\alpha}\left(\sum_{i=1}^{2} h_i x_i^\pi(0)\right) + \mathbb{E}\left[\int_0^t e^{-\alpha s}\sum_{i=1}^{2} h_i A_i(s)ds\right] \\ & \mathbb{E}\left[\int_0^t e^{-\alpha s}c(a^\pi(s))ds\right] + \mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL^\pi_{12}(s) + R_{21}dL^\pi_{21}(s)\Big)\right] - \mathbb{E}\left[\int_0^t e^{-\alpha s}\theta_\pi(s)ds\right]. \end{aligned} \tag{4.20}$$

Proof. Recall that $\{A_i(t); t \geq 0\}, i = 1,2$ denote independent Poisson processes with rates λ_i. Similarly, let $\{D_i(t); t \geq 0\}, i = 1,2$ denote cumulative departure processes corresponding to independent Poisson processes with rates μ_i. Denote the history of these processes up to time t by $\mathcal{F}_t$. The queue dynamics for station i under policy π are,

$$x_i^\pi(s) = x_i^\pi(0) + A_i(s) - D_i\Big(\int_0^s a^\pi(y)\delta_i^\pi(y)dy\Big) \tag{4.21}$$

Substituting (4.21) into the expression for the finite horizon discounted expected cost (4.14) we get,

$$v_{t,\alpha}^\pi = \mathbb{E}\left[\int_0^t e^{-\alpha s}[c(a^\pi(s)) + \sum_{i=1}^2 h_i x_i^\pi(s)]ds\right] + \mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^\pi(s) + R_{21}dL_{21}^\pi\Big)\right. \tag{4.22}$$

$$= \mathbb{E}\left[\int_0^t e^{-\alpha s}c(a^\pi(s))ds\right] + \frac{1-e^{-\alpha t}}{\alpha}\Big(\sum_{i=1}^2 h_i x_i^\pi(0)\Big) + \mathbb{E}\left[\int_0^t e^{-\alpha s}\sum_{i=1}^2 h_i A_i(s)ds\right]$$
$$- \mathbb{E}\Big[\int_0^t e^{-\alpha s}\sum_{i=1}^2 h_i D_i\Big(\int_0^s a^\pi(y)\delta_i^\pi(y)dy\Big)ds\Big] + \mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^\pi(s) + R_{21}dL_{21}^\pi(s)\Big)\right]$$

$$= \mathbb{E}\left[\int_0^t e^{-\alpha s}c(a^\pi(s))ds\right] + \frac{1-e^{-\alpha t}}{\alpha}\Big(\sum_{i=1}^2 h_i x_i^\pi(0)\Big) + \mathbb{E}\left[\int_0^t e^{-\alpha s}\sum_{i=1}^2 h_i A_i(s)ds\right]$$
$$- \mathbb{E}\Big[\int_0^t e^{-\alpha s}\sum_{i=1}^2 h_i \mu_i\Big(\int_0^s a^\pi(y)\delta_i^\pi(y)dy\Big)ds\Big] + \mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^\pi(s) + R_{21}dL_{21}^\pi(s)\Big)\right],$$

where the last equality follows since the process $\{a^\pi(t)\delta_i^\pi(t), t \geq 0\}$ is non-decreasing in t a.s. and adapted to $\mathcal{F}_t$, therefore $\mathbb{E}\Big[D_i\Big(\int_0^s a^\pi(y)\delta_i^\pi(y)dy\Big)\Big] = \mathbb{E}\Big[\mathbb{E}\Big[D_i\Big(\int_0^s a^\pi(y)\delta_i^\pi(y)dy\Big)|\mathcal{F}_s\Big]\Big] = \mathbb{E}[\mu_i \int_0^s a^\pi(y)\delta_i^\pi(y)dy]$. The result follows from the definition of $\theta_\pi(s)$. ∎

Lemma 4.4.2 shows that out optimization criteria is equivalent to maximizing

$$J^{\pi}_{t,\alpha} = \mathbb{E}\left[\int_0^t e^{-\alpha s}\theta_{\pi}(s)ds\right] - \mathbb{E}\left[\int_0^t e^{-\alpha s}c(a^{\pi}(s))ds\right] - \mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL^{\pi}_{12}(s) + R_{21}dL^{\pi}_{21}(s)\Big)\right] \tag{4.23}$$

We are now ready to show the main structural result of this section, which proves that an exhaustive policy at station 1 maximizes $J^{\pi}_{t,\alpha}$.

Theorem 4.4.3. *There exists an optimal policy that is exhaustive at station 1.*

Proof. Let $t > 0$ and consider any admissible policy $\pi_1 \in \Pi$ that is not exhaustive at station 1. Let $0 < \sigma < t$ be the first time such that the servers are set-up at station 1 and π_1 decides to switch to station 2. Let τ be the first time after σ at which the policy π_1 leaves station 2. Also let ν be the time after σ when the policy π_1 finishes the service of the job at the head of station 1 for the first time.

We will construct two processes on the same probability space $(\Omega, \mathcal{B}, \mathcal{P})$ so both the processes see the same arrival and departure sequence. Process 1 follows policy π_1 and process 2 follows policy π_2 defined next. Policy π_2 uses the same amount of resources as policy π_1 at all times i.e, $a_{\pi_1}(s) = a_{\pi_2}(s) = a(s)$ for all $s \geq 0$ almost surely. Furthermore, in terms of scheduling, policy π_2 follows policy π_1 until time σ and after time ν. At time σ, it serves station 1 until the service of the customer here gets completed or time τ whichever occurs first. Denote this time by σ^*. After time σ^*, π_2 follows π_1 except that π_2 serves station 2 if π_1 is serving station 1 in process 1 and the job at the head of station i at time σ has already finished service under process 2. Let $[\zeta_n, \zeta_{n'}]$ denote the n^{th} interval in (σ^*, ν) of total say N intervals during which process 1 serves at station 1 and process 2 serves at station 2. Note that at time ν, π_1 is at station 1 so π_2 also switches to 1 (if not already there) at time ν. The result is that the number of jobs at both stations and stations for which servers are set-up are the same under both processes between time $[0, \sigma) \bigcup [\nu, \infty)$. Figures 4.4(a) and (b) show this scheme for the cases when $\sigma^* < \tau$ and $\sigma^* = \tau$ respectively.

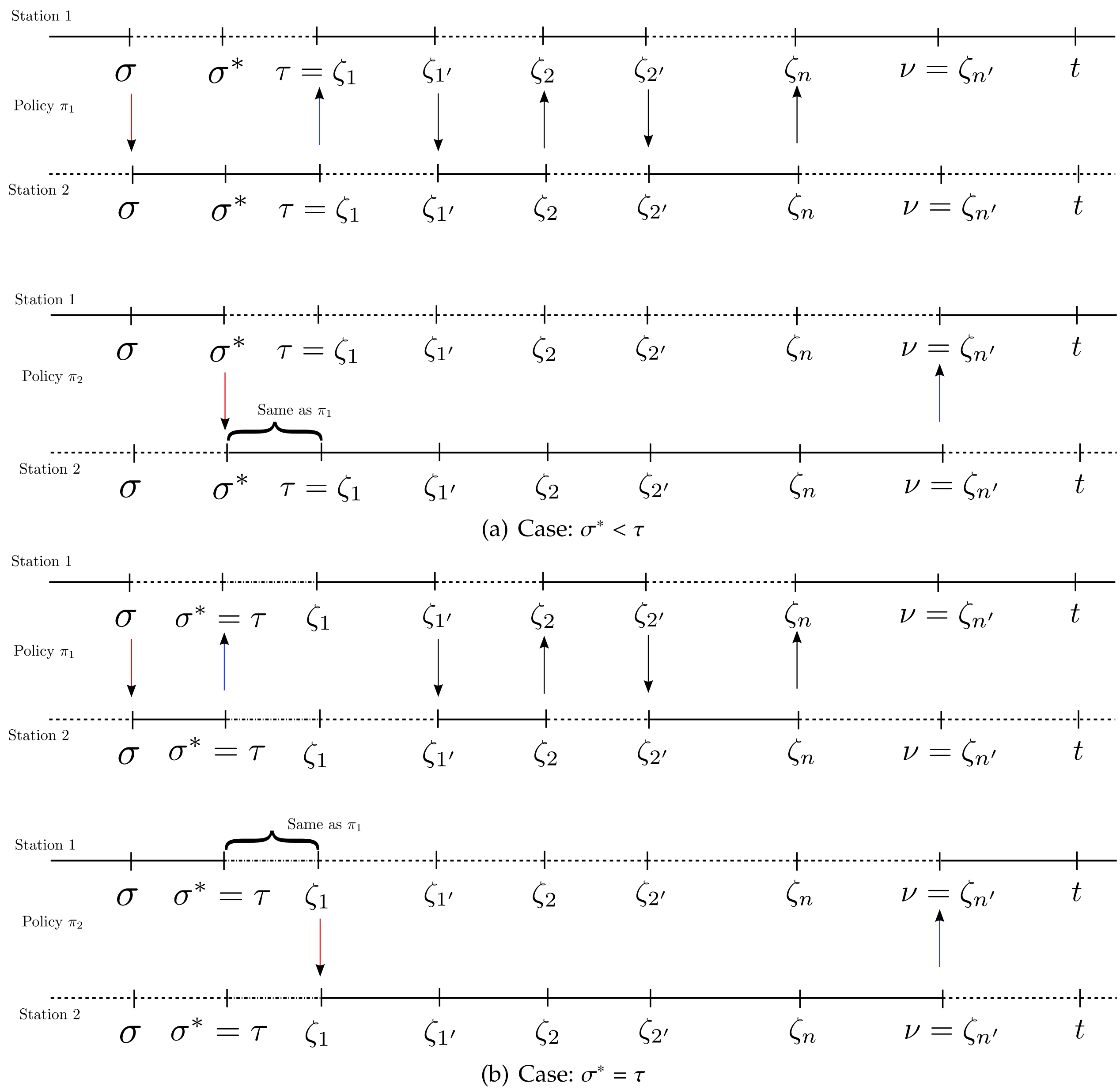


Figure 4.3: Description of policies π_1 and π_2 over time period $[\sigma, \nu]$ for proof of Theorem 4.4.3

.

Since both processes finish the service of exactly one customer at station 1,

$$\sum_{n=1}^{N} \int_{\zeta_n}^{\zeta_n'} a(s)ds = \int_{\sigma}^{\sigma^*} a(s)ds, \tag{4.24}$$

$$\delta_1^{\pi_2}(s) - \delta_1^{\pi_1}(s) = \begin{cases} 0, & s \in [0,\sigma) \bigcup [\sigma^*, \nu] \setminus \bigcup_{k=1}^{n} [\zeta_k, \zeta_{k'}]; \\ 1, & s \in [\sigma, \sigma^*]; \\ -1, & s \in \bigcup_{k=1}^{n} [\zeta_k, \zeta_{k'}], \end{cases} \tag{4.25}$$

and

$$\delta_2^{\pi_2}(s) - \delta_2^{\pi_1}(s) = \begin{cases} 0, & s \in [0,\sigma) \bigcup [\sigma^*, \nu] \setminus \bigcup_{k=1}^{n} [\zeta_k, \zeta_{k'}]; \\ -1, & s \in [\sigma, \sigma^*]; \\ 1, & s \in \bigcup_{k=1}^{n} [\zeta_k, \zeta_{k'}]. \end{cases} \tag{4.26}$$

Consequently we have

$$\theta_{\pi_2}(s) - \theta_{\pi_1}(s) = \begin{cases} 0, & s < \sigma; \\ (\mu_1 h_1 - \mu_2 h_2) \int_{\sigma}^{s} a(y)dy, & \sigma \le s \le \sigma^*, \\ (\mu_1 h_1 - \mu_2 h_2) \int_{\sigma}^{\sigma^*} a(y)dy, & \sigma^* \le s \le \zeta_1, \\ (\mu_1 h_1 - \mu_2 h_2) \Big[\int_{\sigma}^{\sigma^*} a(y)dy - \sum_{k=1}^{\ell-1} \int_{\zeta_k}^{\zeta_{k'}} a(y)dy - \int_{\zeta_\ell}^{s} a(y)dy \Big], & \zeta_\ell \le s \le \zeta_{\ell'}, \\ (\mu_1 h_1 - \mu_2 h_2) \Big[\int_{\sigma}^{\sigma^*} a(y)dy - \sum_{k=1}^{\ell} \int_{\zeta_k}^{\zeta_{k'}} a(y)dy \Big], & \zeta_{\ell'} < s \le \zeta_{\ell+1}, \\ (\mu_1 h_1 - \mu_2 h_2) \Big[\int_{\sigma}^{\sigma^*} a(y)dy - \sum_{k=1}^{n} \int_{\zeta_k}^{\zeta_{k'}} a(y)dy \Big], & s > \zeta_n. \end{cases} \tag{4.27}$$

Using the fact $\mu_1 h_1 \ge \mu_2 h_2$ and (4.24) in (4.27) now implies that $\theta_{\pi_2}(s) \ge \theta_{\pi_1}(s)$ for all $s \in [0, t]$. Thus

$$\int_0^t e^{-\alpha s} \theta_{\pi_1}(s) ds \le \int_0^t e^{-\alpha s} \theta_{\pi_2}(s) ds \tag{4.28}$$

We next look at the difference in switching costs accrued under the two policies. This is crucially dependent on two specific switches, the first one from 1 to 2 and the second switch from 2 to 1, that policy π_2 takes at a later time than policy π_1. The first switch is the one made by π_1 from station 1 to 2 at time σ, policy π_2 makes this switch at σ^* or later but before time ν, we denote this time by t_{12}. These switches are denoted by red arrows in Figure 4.4. The second switch is made by policy π_1 from station 2 to 1 at time τ, the corresponding switch under policy π_2 comes at time ν. These switches are denoted by blue arrows in Figure 4.4. Rest of the switches are made at the same time under both policies although π_1 may make more switches than π_2 (e.g, at times ζ_n and $\zeta_{n'}$ for $n > 1$). Consequently, the total switching cost incurred until time t, under policy π_1 is greater than that incurred under π_2 almost surely. In particular, the difference is the switching cost under the two policies for any $\alpha \geq 0$ is

$$\int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^{\pi_1}(s) + R_{21}dL_{21}^{\pi_1}(s)\Big) - \int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^{\pi_2}(s) + R_{21}dL_{21}^{\pi_2}(s)\Big) \tag{4.29}$$
$$\geq R_{12}e^{-\alpha\sigma} + R_{21}e^{-\alpha\tau} - \Big(R_{12}e^{-\alpha t_{12}} + R_{21}e^{-\alpha\nu}\Big) \geq 0,$$

where the first inequality is due to ignoring additional switches that π_1 may make at times $\{\zeta_n\}$ and $\{\zeta_{n'}\}$ and the second inequality is because $t_{12} > \sigma$ and $\nu > \tau$.

Let σ_1 be the time at which π_2 violates the exhaustive scheduling rule. Note that $\sigma_1 > \sigma$. Using the same argument as before we can construct a sequence of policies $\{\pi_m; m = 1, 2, \ldots\}$ such that $\theta_{\pi_m}(s) \geq \theta_{\pi_{m-1}}(s)$ for all $s \in [0, t]$ with $\sigma_m > \sigma_{m-1}$. Moreover the switching cost under π_m is lower than that under π_1. Thus there exists a policy π_n such that $\theta_{\pi_n}(s) \geq \theta_{\pi_{n-1}}(s) \ldots \geq \theta_{\pi_1}(s)$ for all $s \in [0, t]$ with $\sigma_{n-1} \leq t \leq \sigma_n$. Let $\pi = \pi_n$, then π

- is non-anticipative,

- serves at station 1 in exhaustive manner,
- saves more on holding cost than π_1, i.e, $\theta_\pi(s) \geq \theta_{\pi_1}(s)$ for all $s \in [0,t]$, a.s therefore

$$\mathbb{E}\left[\int_0^t e^{-\alpha s}\theta_{\pi_1}(s)ds\right] \leq \mathbb{E}\left[\int_0^t e^{-\alpha s}\theta_{\pi}(s)ds\right]$$

- incurs lesser set-up cost than π_1 i.e,

$$\mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^{\pi_1}(s) + R_{21}dL_{21}^{\pi_1}(s)\Big) \geq \mathbb{E}\left[\int_0^t e^{-\alpha s}\Big(R_{12}dL_{12}^{\pi}(s) + R_{21}dL_{21}^{\pi}(s)\Big)\right]$$

- uses same amount of resources as π_1 at any time i.e, $a_\pi(s) = a_{\pi_1}(s) = a(s)$ a.s. for all $s \geq 0$, therefore

$$\mathbb{E}\left[\int_0^t e^{-\alpha s}c(a^\pi(s))ds\right] = \mathbb{E}\left[\int_0^t e^{-\alpha s}c(a_1^\pi(s))ds\right]$$

Thus $J_{t,\alpha}^\pi \geq J_{t,\alpha}^{\pi_1}$ for all $\alpha \geq 0$, where $J_{t,\alpha}^\pi$ is defined in (4.23). Since t is arbitrary, the result holds for the infinite horizon case. Hence the exhaustive scheduling rule is optimal in Π for both long-run discounted cost and average cost case. ■

Theorem 4.4.3 provides some insights into the nature of the optimal policy, the servers switch from station 1 to 2 only when station 1 is empty. However it is still not clear as to when to switch when the servers are at station 2 nor if the number of servers being used are non-decreasing in queue length. Despite our best efforts, we were not able to show theoretical results to answer these questions. Instead, we present a numerical study in the next section to complete this picture.

4.5 Numerical Study and Heuristics

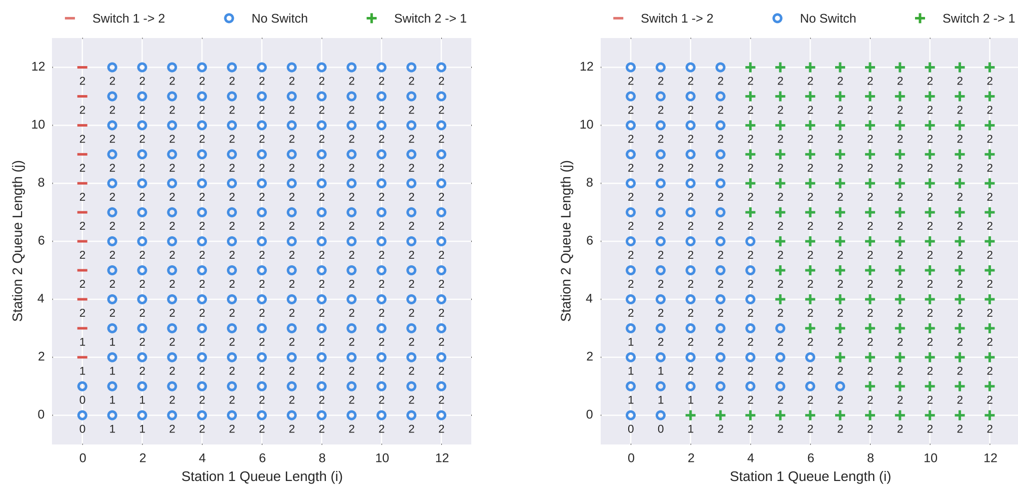
In this section we conduct a numerical study on the Consolidated Switching (CS) model. We use Example 4.5.1 to augment the theoretical insights in to the nature of optimal policy that was provided in the previous section. We discover several interesting and perhaps surprising features of the optimal policy for the consolidated switching model. In Example 4.5.2 we undertake a study to compare the performance of the consolidated switching model with respect to the model with unrestricted switching. To make the intuition clear and keep the exposition simple, both problems deal with a system with two stations. In all the examples discussed here, we used policy iteration method for solving the average cost problem and truncated the queue lengths to 40.

Example 4.5.1. We consider the same set-up as in Example 4.2.1. The optimal average cost policy under the consolidated switching regime for this example is shown in Figure 4.4 for various queue lengths and server set-up configurations.

Based on Figure 4.4, we make several important observations about the scheduling policy:

Observation 1 - Exhaustive scheduling and threshold switching policy at the high cost station: As predicted by Theorem 4.4.3, Figure 4.4(a) shows that when servers are set-up for station 1, they are not switched to station 2 until station 1 becomes empty ($x_1 = 0$). Moreover, when station 1 becomes empty, servers switch after the queue-length at station 2 exceeds a certain threshold (2 jobs in this example).

Observation 2 - Switching server from the low cost station to the high cost station follows a switching curve but it is not monotone: Figure 4.4(b) shows that when the



(a) Optimal policy if servers are set-up at station 1

(b) Optimal policy if servers are set-up at station 2

Figure 4.4: Optimal policy for Example 4.5.1 (CS Model). Symbols represent if servers are switching away (-), towards (+) or staying (o) aet station 1. The number below each symbol denotes the number of active servers.

servers are set-up at station 2, it is not always optimal to serve there exhaustively. If there are sufficiently many jobs at station 1, it may be optimal to switch to 1 even when there are pending jobs at station 2. The intuition behind this is that by working at station 1, the holding costs can be reduced at a higher rate than by working at 2; this may justify paying the additional switching cost. However, the switching curve is not monotone in j and displays complex characteristics.

Observation 3 - Number of servers active at each station are non-decreasing in queue length: As seen for the case without set-up costs, the number of servers working are non-decreasing in the queue length of each station.

The next example shows the effect of switching cost on the performance of the consolidated switching policy.

Example 4.5.2. Consider a system with two stations ($m = 2$). The total service capacity available to the system is, $S = 3$ and the cost of using s resources is $c(s) = s^3$. The arrival rates are $\lambda_1 = \lambda_2 = 0.1$ and the service rates are $\mu_1 = \mu_2 = 0.09$. The system is operating under medium load condition with total load $(\sum_{i=1}^{m} \frac{\lambda_i}{S\mu_i}) = 0.6$. The holding costs are $h_1 = 2$ and $h_2 = 1$. The set-up costs are symmetric ($R_{ij} = R_{ji}$). We plot the long-run average cost incurred by the consolidated switching policy and the full model presented in Section 4.2 (labeled optimal) as the total set-up cost ($S \cdot R_{ij}$) varies between 10^{-3} to 10^3. For reference, we also plot the cost corresponding to the Static Rate (SR) and the Gated Static (GS) heuristics presented in Section 3.5. Figure 4.5 shows that the

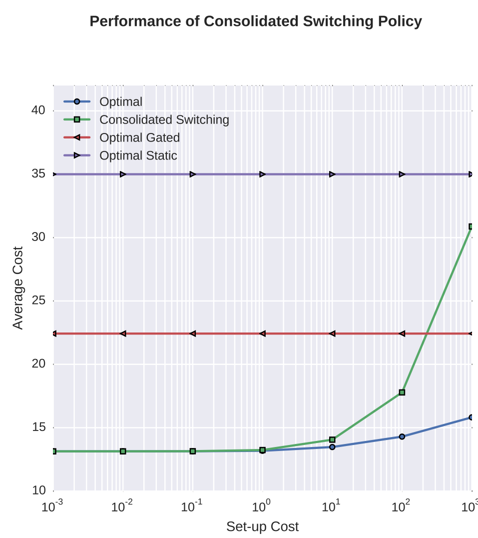


Figure 4.5: Performance of the consolidated switching policy

consolidated switching policy performs well when the total switching costs are less than 10. Its performance starts to deteriorate rapidly as the switching costs increase beyond 10 and becomes worse than the performance of the Gated Static policy for switching costs above 100.

4.5.1 Fluid Heuristics

Example 4.5.2 shows that when the total switching costs are greater than 10, we can't rely on the consolidated switching policy. For such scenarios, we develop an alternate heuristic by treating the queues as deterministic fluid queues. Similar heuristics have been considered for the model without the utilization cost by Duenyas et al. [18], [19] and Slegers et al.[51].

Consider the case of two stations ($m = 2$). Suppose there is an arrival at station i. If there are more servers set-up at this station then are currently working i.e, $u_i > s_i$ and if all stations are currently stable ($\lambda_i < s_i\mu_i$), then we compare the action of turning on one server vs not doing anything. If we don't do anything

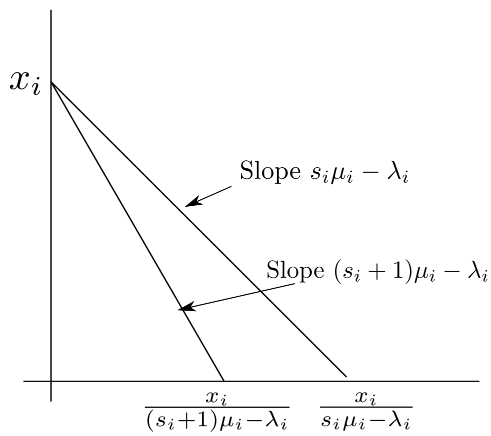


Figure 4.6: Evolution of the Deterministic Fluid Queue

and suppose we use s_i servers until the fluid queue gets empty then station i will get empty in $\frac{x_i}{s_i\mu_i-\lambda_i}$ units of time (see Figure 4.6) and the system will incur the following cost upto that time

$$c = c(s_1 + s_2)\frac{x_i}{s_i\mu_i - \lambda_i} + \frac{h_i x_i^2}{2(s_i\mu_i - \lambda_i)} + \text{ holding cost at the other station.}$$

While if we turn on an additional server at station i, we incur

$$c^{+} = c(s_1+s_2+1)\frac{x_i}{(s_i+1)\mu_i - \lambda_i} + \frac{h_i x_i^2}{2((s_i+1)\mu_i - \lambda_i)} + \text{holding cost at the other station.}$$

Thus, we will turn on a server at station i if $c^{+} < c$. If all the servers at station i are busy i.e, $s_i = u_i$, then we consider the action of switching one server from the other station (j) to station i (assuming that station j will remain stable after this switch), in this case the cost incurred until both stations get empty is

$$c^{\rightarrow} = c(s_1+s_2)\max\left\{\frac{x_i}{(s_i+1)\mu_i}, \frac{x_j}{(s_j-1)\mu_j - \lambda_j}\right\} + \frac{h_i x_i^2}{2((s_i+1)\mu_i - \lambda_i)} + \frac{h_j x_j^2}{2((s_j-1)\mu_j - \lambda_j)} + R_{ji}.$$

While if we don't switch then the cost incurred until both stations are empty is

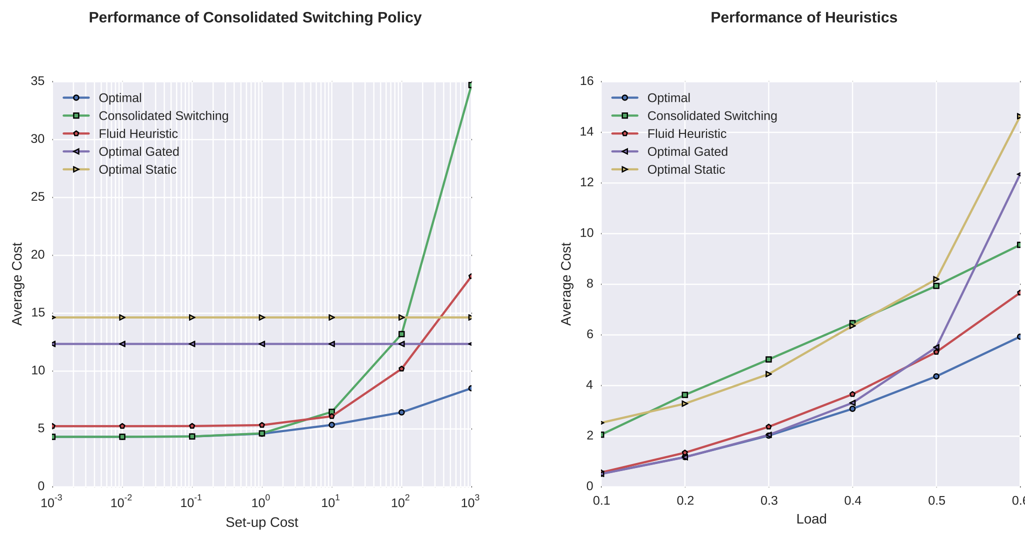
$$c = c(s_1+s_2)\max\left\{\frac{x_i}{s_i\mu_i - \lambda_i}, \frac{x_j}{s_j\mu_j - \lambda_j}\right\} + \frac{h_i x_i^2}{2(s_i\mu_i - \lambda_i)} + \frac{h_j x_j^2}{2(s_j\mu_j - \lambda_j)}.$$

So we switch a server from station i to station j if $c^{\rightarrow} < c$.

If there is departure of a job from station i, we use a similar analysis and compare the action sequences to decide whether to do nothing, turn-off a server or switch a server to the other station.

We compare the performance of the fluid heuristic using a simulation study in the next example.

Example 4.5.3. Consider a system with two stations ($m = 2$). The total service capacity available to the system is, $S = 3$ and the cost of using s resources is $c(s) = \frac{s^2}{2}$. The arrival rates are $\lambda_1 = \lambda_2 = 0.1$ and the service rates are $\mu_1 = \mu_2 = 0.09$. The system is operating under medium load condition with total load ($\sum_{i=1}^{m} \frac{\lambda_i}{S\mu_i}$) = 0.6. The holding costs are $h_1 = 2$ and $h_2 = 1$. The set-up costs are symmetric ($R_{ij} = R_{ji}$). We compare the performance of several heuristics as the total set-up cost varies between 10^{-3} to 10^{3} in Figure 4.7(a). We see that when set-up costs less than 10, the Consolidate Switching (CS) policy performs



(a) Average Cost vs Total Set-up Cost (b) Average Cost vs Load

Figure 4.7: Comparison of Heuristic Policies

slightly better than the fluid heuristic. However for set-up costs between 10 and 100, the fluid heuristic performs significantly better than the CS policy. Figure 4.7(b) shows a comparison of the performance of heuristics as the load varies between 0.2 and 0.6 for a fixed set-up cost $R_{ij} = 40$. In this case we see that when the load is low (less than 0.3), the Gated Static heuristic performs the best. For medium to high loads the performance of the Gated Static heuristic starts to degrade rapidly while the fluid heuristic fairs relatively better for loads above 0.5.

4.6 Conclusions

In this Chapter we investigated the model with multiple job types in the presence of switching costs. Assuming that the goal is to minimize a combination of effort cost, holding cost and switching costs, we study this problem under both the discounted and average cost optimality criterion. We showed that the optimal policy is shows many peculiarities that are hard to characterize. When the switching costs are low, a model where servers switch is a consolidated fashion provides a tractable heuristic for the original model. In this case we partially characterized the optimal policy as being exhaustive at the highest priority (based on the $c - \mu$ rule) queue. Using numerical studies, we show that the performance of the consolidated switching model suffers considerably for medium to high switching costs. For such scenarios, we develop a heuristic for the original model based on deterministic fluid queues. Our numerical study shows that this heuristic provides good performance for medium switching costs and when the load is in medium to high range.

This work can be extended in several directions. Firstly, adding server switch-over times to this model will be beneficial from the practical viewpoint. In this case exact analysis using MDP framework will prove challenging. Instead, focussing on finding good approximation schemes for such systems will be more fruitful. Secondly, adding time-varying arrival rates will also add to the practical utility of the model. A framework based on the Markov modulated Poisson processes like the one presented in Chapter 1 may provide a tractable avenue for modeling such time-varying rates.

APPENDIX A

APPENDIX FOR CHAPTER 2

This appendix is dedicated to providing proofs of Propositions 2.2.1 and 2.2.2 from Chapter 2. The results of Dai [15] and Dai and Meyn [16] are utilized to show the stability of a stochastic model by establishing the stability of its fluid limit approximation. For the purpose of this analysis, consider the continuous time Markov process, $X^{\pi}(t) = \{(Q^{\pi}(t), S^{\pi}(t)), t \geq 0\}$, induced by an admissible stationary policy $\pi \in \Pi$ where $\{Q^{\pi}(t), t \geq 0\}$ and $\{S^{\pi}(t), t \geq 0\}$ represent the queue length and phase transition process for arrivals, respectively.

The proof approach for Proposition 2.2.1 follows closely that of Kaufman and Lewis [31] (Proposition 3.1) and is done in several steps. Let $\hat{\pi}$ be a policy that selects a constant rate $\hat{\mu} \in (\sum_{s=1}^{L} p_s \lambda_s, \bar{u}]$ whenever the queue is non-empty. Let $X^{\hat{\pi}}(t) = \{(Q^{\hat{\pi}}(t), S^{\hat{\pi}}(t)), t \geq 0\}$ be the Markov process induced by the policy $\hat{\pi}$ on state space $\mathbb{X} = \{(n, s) \mid n \in \mathbb{Z}^{+}, s \in \mathcal{S}\}$. Since the policy is fixed for the remainder of this section, in the interest of brevity we suppress dependence on $\hat{\pi}$. The norm of a state, $x = (n, s) \in \mathbb{X}$, is defined to be $|x| := n + s$. For an initial state $X(0) = x$, we define the scaled queue length process

$$\bar{Q}^{x}(t) := \frac{1}{|x|} Q^{x}(|x|t).$$

We will use a similar notation to denote scaled versions of other stochastic processes.

For each $s \in \mathcal{S}$, let $\{\xi_s(k), k \geq 1\}$ be a sequence of i.i.d exponential random variables with mean $1/\lambda_s$. The sequence $\{\xi_s(k), k \geq 1\}$ represents the set of job inter-arrival times when the arrival process is in phase s. Also, let $\{\eta(k), k \geq 1\}$ be a sequence of i.i.d exponential random variables with mean 1 representing the set of job completion times. Based on these sequences, we define the following cumulative processes

$$E_s(t) = \max\{k \geq 0 \mid \xi_s(1) + \xi_s(2), \ldots, \xi_s(k) \leq t\} \qquad \text{for } s \in \mathcal{S},$$
$$D(t) = \max\{k \geq 0 \mid \eta(1) + \eta(2), \ldots, \eta(k) \leq t\}.$$

Let $Y_s^x(t)$ be the cumulative amount of time the arrival process spends in phase s until time t when the initial state is x. Similarly, let $T^x(t)$ be the cumulative amount of time for which there is at least one customer in the queue, $I^x(t)$ be the cumulative amount of time when the queue is empty and $W^x(t)$ be the total work done by the server until time t. We can now write the following system of equations for the stochastic process induced by policy $\hat{\pi}$ when starting from initial state x,

$$Q^x(t) = Q^x(0) + \sum_{s=1}^{L} E_s(Y_s^x(t)) - D(W^x(t)), \tag{A.1}$$

$$Q^x(t) \geq 0, \tag{A.2}$$

$$\sum_{s=1}^{L} Y_s^x(t) = t, \tag{A.3}$$

$$W^x(t) = \hat{\mu} T^x(t), \tag{A.4}$$

$$T^x(t) + I^x(t) = t, \tag{A.5}$$

$$\int_0^{\infty} Q^x(t) dI^x(t) = 0, \tag{A.6}$$

$$Y_s^x(t), T^x(t), I^x(t), W^x(t) \quad \text{start from zero and are non-decreasing in } t. \tag{A.7}$$

A few comments are in order. First, note that (A.6) imposes the constraint that the server is idle only when the system is empty. For a subsequence $\{x_n, n \geq 1\}$ such that $|x_n| \to \infty$, any limit point, $\bar{Q}(t)$, of the sequence $\{\bar{Q}^{x_n}, n \geq 1\}$ is called a *fluid limit*. It will be shown that every fluid limit satisfies a set of equations known as the *fluid model*. A fluid model is called *stable* if there exists a $t_0 > 0$

such that $\bar{Q}(t) = 0$ for all $t \geq t_0$ and for all fluid limits. We next present the fluid model and convergence results for the scaled processes

Proposition A.0.1. *Let $\{x_j \mid x_j \in \mathbb{X}, j \geq 1\}$ be a sequence of initial states with $|x_j| \to \infty$. Then with probability 1, there exists a subsequence, $\{x_{j_k}, k \geq 1\}$, such that*

$$(\bar{Q}^{x_{j_k}}(0), \bar{S}^{x_{j_k}}(0)) \to (\bar{Q}(0), 0), \tag{A.8}$$

$$(\bar{Q}^{x_{j_k}}(t), \bar{T}^{x_{j_k}}(t)) \to (\bar{Q}(t), \bar{T}(t)) \quad \textit{uniformly on compact sets (u.o.c.)}, \tag{A.9}$$

where $(\bar{Q}(t), \bar{T}(t))$ satisfy the following equations,

$$\bar{Q}(t) = \bar{Q}(0) + \sum_{s=1}^{L} p_s \lambda_s t - \bar{W}(t), \tag{A.10}$$

$$\bar{Q}(t) \geq 0, \tag{A.11}$$

$$\bar{W}(t) = \hat{\mu}\bar{T}(t), \tag{A.12}$$

$$\bar{T}(t) + \bar{I}(t) = t, \tag{A.13}$$

$$\int_0^\infty \bar{Q}(t) d\bar{I}(t) = 0, \tag{A.14}$$

$$\bar{T}(t), \bar{I}(t), \bar{W}(t) \quad \textit{start from zero and are non-decreasing in } t. \tag{A.15}$$

Proof. Since $\bar{Q}^{x_j}(0) \leq 1$, $\bar{S}^{x_j}(0) \leq 1$ and $1 \leq S^{x_j}(0) \leq L$ for all $j \in \mathbb{N}$, there exists a subsequence, $\{x_{j_k}, k \geq 1\}$ such that $(\bar{Q}^{x_{j_k}}(0), \bar{S}^{x_{j_k}}(0)) \to (\bar{Q}(0), 0)$. For any fixed sample path ω and $0 \leq s \leq t$, we have $0 \leq \bar{T}^{x_j}(t) - \bar{T}^{x_j}(s) \leq t - s$. Thus, the function $\bar{T}^{x_j}(t)$ is uniformly Lipschitz of order 1. Since $0 \leq \bar{T}^{x_j}(t) \leq t$, it is also uniformly bounded for each $j \geq 1$. Therefore, the sequence $\{\bar{T}^{x_j}(t), j \geq 1\}$ is equicontinuous. By Arzela-Ascoli theorem, any subsequence of $\{\bar{T}^{x_j}(t), j \geq 1\}$ has a u.o.c. convergent subsequence.

Since the phase transition process is ergodic, for each $s \in \{1, \ldots, L\}$ we have with probability 1 $\lim_{t\to\infty} Y_s^x(t)/t = p_s$. Furthermore, from the strong law of large numbers for renewal processes, the following hold almost surely

$$\lim_{t\to\infty} E_s(t)/t = \lambda_s \quad s \in \mathcal{S},$$

$$\lim_{t\to\infty} D(t)/t = 1.$$

The above results can be used in a manner similar to Lemma 4.2 of [15], to yield (with probability 1)

$$\bar{Y}_s(t) = \lim_{k\to\infty} \frac{1}{|x_{j_k}|} Y^{x_{j_k}}(|x_{j_k}|t) = p_s t \quad u.o.c., \text{ for } s \in \mathcal{S}, \tag{A.16}$$

$$\bar{E}_s(t) = \lim_{k\to\infty} \frac{1}{|x_{j_k}|} E(|x_{j_k}|t) = \lambda_s t \quad u.o.c., \text{ for } s \in \mathcal{S} \tag{A.17}$$

$$\bar{D}(t) = \lim_{k\to\infty} \frac{1}{|x_{j_k}|} D(|x_{j_k}|t) = t \quad u.o.c. \tag{A.18}$$

The equality in (A.10) follows from (A.1) and (A.16)-(A.18). Similarly, (A.12)-(A.15) follow directly from (A.4)-(A.7), respectively. ∎

Proof of Proposition 2.2.1: We start by showing that the fluid model provided in (A.10)-(A.15) is stable. First note that $\bar{T}(t), \bar{I}(t)$ and $\bar{Y}_s(t)$ are Lipschitz continuous and therefore absolutely continuous and differentiable almost everywhere. Taking the derivative with respect to t in (A.10) and (A.12) yields

$$\dot{\bar{Q}}(t) = \sum_{s=1}^{L} p_s\lambda_s - \hat{\mu}\dot{\bar{T}}(t).$$

Further, due to the non-idling constraint (A.14), $\dot{\bar{I}}(t) = 0$ whenever $\bar{Q}(t) > 0$. Thus, from (A.13), $\dot{\bar{T}}(t) = 1$ whenever $\bar{Q}(t) > 0$. So for $\bar{Q}(t) > 0$, we have

$$\dot{\bar{Q}}(t) = \sum_{s=1}^{L} p_s\lambda_s - \hat{\mu}.$$

Our choice of the stationary policy enabled by the stability condition (2.6), implies that $\dot{\bar{Q}}(t) < 0$ whenever $\bar{Q}(t) > 0$. Thus from Lemma 5.2 of [15] we have that the fluid limit process for queue length is non-increasing and there exists a $t_0 \geq 0$ such that $\bar{Q}(t) = 0$ for all $t \geq t_0$. That is, the fluid model is stable. The results of Theorem 4.2 of [15] imply that the Markov process induced by the stationary policy $\hat{\pi}$, is positive recurrent and stationary distribution exists. Furthermore, since the embedded discrete time Markov chain for the process is irreducible, this process is *ergodic*. This implies that the long-run average cost under $\hat{\pi}$, say $g^{\hat{\pi}}$, is independent of the initial state.

To show that $g^{\hat{\pi}}$ is finite, we use the results of Theorem 4.1(i) of [16]. Since the fluid model is stable and conditions A1) and A2) of Theorem 4.1 hold, it follows that for any integer $p \geq 1$ $\limsup_{t\to\infty} \frac{1}{t} \int_0^t \mathbb{E}_x[|Q(u)|^p]du < \infty$ for each initial condition x. Since the holding cost has a polynomial rate of growth, we have that the long-run average holding cost rate is finite. Moreover, the direct contribution to long-run cost rate due to serving at $\hat{\mu}$ whenever the queue is not empty is at most $c(\hat{\mu}) < \infty$. It therefore follows $g^{\hat{\pi}}$ is finite.

It remains to consider the necessity of (2.6). Consider the Markov process induced by a policy that uses the highest available service rate whenever the queue is not empty. As shown by Yechiali[57] using the detailed balance equations for the steady state distribution a non-trivial invariant measure exists only if $\bar{u} > \sum_{s=1}^{L} p_s \lambda_s$. The result follows and the proof is complete. ∎

The remainder of this section is dedicated to proving Proposition 2.2.2 holds. We show that the optimal value and relative value functions satisfying the ACOI, (2.5), exist and can be obtained via limits from the discounted expected cost value functions. Thus, the structural results proved for the discounted cost case continue to hold for the average cost case. In proving these results, we verify the following set of assumptions (SEN) (included for completeness) provided by Sennott [48] hold.

- *SEN1:* There exist $\delta > 0$ and $\epsilon > 0$ such that for every state and action, there is a probability of at least ϵ that the transition time will be greater that δ.
- *SEN2:* There exists B such that $\tau(i, \mu) \leq B$ for every state i and control μ, where $\tau(i, \mu)$ is the expected transition time out of state i when control μ is chosen.
- *SEN3:* $v_\alpha(i) < \infty$ for all states i and $\alpha > 0$.
- *SEN4:* There exists $\alpha_0 > 0$ and nonnegative numbers M_i such that $w_\alpha(i) \leq M_i$ for every state i and $0 < \alpha < \alpha_0$ where $w_\alpha(i) = v_\alpha(i) - v_\alpha(\mathbf{0})$, for distinguished state $\mathbf{0}$. For every state i, there exists an action μ_i such that $\sum_j P_{ij}(\mu_i) M_j < \infty$.
- *SEN5:* There exists $\alpha_0 > 0$ and a non-negative number N such that $-N \leq w_\alpha(i)$ for every i and $0 \leq \alpha \leq \alpha_0$.
- *SEN6:* For each state i, the expected single stage discounted cost $f_\alpha(i, \mu)$ is a lower semi-continuous (lsc) function on the product space $[0, \infty) \times \mathbb{A}$. Note that $f_0(i, \mu)$ is the un-discounted single stage cost.
- *SEN7:* For all states i and j, the function $L_{ij}(\alpha, \mu) = P_{ij}(\mu) \int_{t=0}^{\infty} e^{-\alpha t} \nu e^{-\nu t} dt$ is a lsc function on the product space $[0, \infty) \times \mathbb{A}$.
- *SEN8:* Assume that α_n is a sequence of discount factors converging to 0 with the property that π_{α_n}, the associated sequence of α-discount optimal policies, converge to a stationary policy π. Then for each state i, $\liminf_n \tau(i, \pi_{\alpha_n}) \leq \tau(i, \pi)$.

Proof of Proposition 2.2.2: We begin by verifying the *SEN* assumptions. Since the uniformizing rate is strictly positive and finite, assumptions *SEN1* and *SEN2* hold. It was shown in Proposition 2.2.1 that under the stability condition (2.6), there exists a stationary policy that induces an ergodic Markov process with finite long-run average expected cost. Thus, the hypotheses of Lemma 2 of [48] hold; validating *SEN3 and SEN4* assumptions. Let $\bar{s} = \arg\min_{s \in \mathcal{S}}\{v_\alpha(0, s)\}$ and define the distinguished state as $\mathbf{0} = (0, \bar{s})$. It follows from Proposition 3.1 that for any $\alpha > 0$, $v_\alpha(\mathbf{0}) \leq v_\alpha(n, s)$ for all $(n, s) \in \mathbb{X}$. Therefore, $w_\alpha(n, s) \geq 0$ and *SEN5* is satisfied. *SEN6* holds since for each $(n, s) \in \mathbb{X}$, $f_\alpha((n, s), \mu) = (c(\mu) + h(n))/(\alpha + \nu)$ is a continuous function on $[0, \infty) \times \mathbb{A}$.

There is no decision to be made when $n = 0$. Fix $n \geq 1$ and $s \in \{1, 2, \ldots, L\}$ and note,

$$L_{(n,s),(n',s')}(\alpha, \mu) = \begin{cases} \frac{\lambda_s}{\alpha+\nu} & \text{if} \quad n' = n + 1, s' = s, \\ \frac{\mu}{\alpha+\nu} & \text{if} \quad n' = n - 1, s' = s, \\ \frac{Q_{ss'}}{\alpha+\nu} & \text{if} \quad n' = n, s' \in \mathcal{S}, \\ 0 & \text{otherwise.} \end{cases} \tag{A.19}$$

So the function $L_{x,x'}(\alpha, \mu)$ is jointly continuous in α and μ for each $x, x' \in \mathbb{X}$, therefore *SEN7* holds. Finally, since for any policy π, $\tau(i, \pi) = 1/\nu$, *SEN8* holds.

SEN1, *SEN3*, *SEN6* and *SEN7* are required in Theorem 11 of [48] to prove the first result. Since the holding costs are assumed to have a polynomial rate of growth, the hypotheses of Proposition 4 of [48] are satisfied. The last two results follow from Theorem 12 (and its proof) of [48]. ∎

APPENDIX B

APPENDIX FOR CHAPTER 3

This Chapter is dedicated to the proofs of Lemmas 3.3.7 and 3.3.14 of Chapter 3.

Proof of Lemma 3.2.1 We may apply Theorem 6.10.4 of [46] to show that $\mathcal{T}^L_\alpha$ is a J stage contraction mapping and $v_{\alpha,k} \uparrow v_\alpha$. In particular let $r(x) := \sum_{i=1}^m h_i x_i + c(S\bar{\mu})$ and $v_r := \sup_{x\in\mathbb{X}} r(x)^{-1}|v(x)|$ be the weighted supremum norm for real-valued functions v on $\mathbb{X}$. Then the cost rate function $C(x,(a,s))$ and transitions probabilities $p((x,(a,s)),y)$ defined in (3.5) satisfy Assumptions 6.10.1 and 6.10.2 of [46] (take $\mu = 1$ for 6.10.1 and $L = \sum_{i=1}^m h_i\lambda_i$ for 6.10.5 (a)).

Poof of Lemma 3.3.7: First note that since property 1 (convexity) follows from properties 2 and 3, we only need to show the last two properties. We use induction and FHOE (3.8) to prove properties 2 and 3. Since $v_{\alpha,0}(i,j) = 0$ for all $i \geq 0, j \geq 0$, the result holds trivially for $k = 0$. For the inductive step, suppose $v_{\alpha,k}(i,j)$ satisfies these properties. We start with property 2. Let $i \geq 2, j \geq 1$ and suppose s^+ and s^- be the optimal resource allocation decision for the $k+1$ stage problem when the state is $(i, j+1)$ and $(i-1, j)$ respectively. Suppose we use the potentially sub-optimal allocation (s_1^-, s_2^+) when the state is $(i-1, j+1)$. The

FHOE (3.8) yield,

$$\begin{aligned}
\Delta^1 v_{\alpha,k+1}(i, j+1) &= (v_{\alpha,k+1}(i, j+1) - v_{\alpha,k+1}(i-1, j+1) \\
&\geq h_1 + \lambda_1 \Delta^1 v_{\alpha,k}(i+1, j+1) + \lambda_2 \Delta^1 v_{\alpha,k}(i, j+2) + c(s_1^+ + s_2^+) - c(s_1^- + s_2^+) \\
&+ (A_1\mu_1 - s_1^+\mu_1) v_{\alpha,k}(i, j+1) + s_1^+\mu_1 v_{\alpha,k}(i-1, j+1) \\
&+ (A_2\mu_2 - s_2^+\mu_2) v_{\alpha,k}(i, j+1) + s_2^+\mu_2 v_{\alpha,k}(i, j) \\
&- (A_1\mu_1 - s_1^-\mu_1) v_{\alpha,k}(i-1, j+1) - s_1^-\mu_1 v_{\alpha,k}(i-2, j+1) \\
&- (A_2\mu_2 - s_2^+\mu_2) v_{\alpha,k}(i-1, j+1) - s_2^+\mu_2 v_{\alpha,k}(i-1, j)
\end{aligned}$$

Similarly using the potentially sub-optimal decision (s_1^+, s_2^-) when the state is (i, j) yields,

$$\begin{aligned}
\Delta^1 v_{\alpha,k+1}(i, j) &= (v_{\alpha,k+1}(i, j) - v_{\alpha,k+1}(i-1, j) \\
&\geq h_1 + \lambda_1 \Delta^1 v_{\alpha,k}(i+1, j) + \lambda_2 \Delta^1 v_{\alpha,k}(i, j+1) + c(s_1^+ + s_2^-) - c(s_1^- + s_2^-) \\
&+ (A_1\mu_1 - s_1^+\mu_1) v_{\alpha,k}(i, j) + s_1^+\mu_1 v_{\alpha,k}(i-1, j) \\
&+ (A_2\mu_2 - s_2^-\mu_2) v_{\alpha,k}(i, j) + s_2^-\mu_2 v_{\alpha,k}(i, j-1) \\
&- (A_1\mu_1 - s_1^-\mu_1) v_{\alpha,k}(i-1, j) - s_1^-\mu_1 v_{\alpha,k}(i-2, j) \\
&- (A_2\mu_2 - s_2^-\mu_2) v_{\alpha,k}(i-1, j) - s_2^-\mu_2 v_{\alpha,k}(i-1, j-1)
\end{aligned}$$

Combining the two inequalities we have,

$$
\begin{aligned}
(\Delta^1 v_{\alpha,k+1}(i,j+1) - \Delta^1 v_{\alpha,k+1}(i,j)) \geq\ & \lambda_1\Big(\Delta^1 v_{\alpha,k}(i+1,j+1) - \Delta^1 v_{\alpha,k}(i+1,j)\Big) \\
& + \lambda_2\Big(\Delta^1 v_{\alpha,k}(i,j+2) - \Delta^1 v_{\alpha,k}(i,j+1)\Big) \\
& + (A_1\mu_1 - s_1^+\mu_1)\Big(\Delta^1 v_{\alpha,k}(i,j+1) - \Delta^1 v_{\alpha,k}(i,j)\Big) \\
& + s_1^-\mu_1\Big(\Delta^1 v_{\alpha,k}(i-1,j+1) - \Delta^1 v_{\alpha,k}(i-1,j)\Big) \\
& + (A_2\mu_2 - s_2^+\mu_2)\Big(\Delta^1 v_{\alpha,k}(i,j+1) - \Delta^1 v_{\alpha,k}(i,j)\Big) \\
& + s_2^-\mu_2\Big(\Delta^2 v_{\alpha,k}(i,j) - \Delta^2 v_{\alpha,k}(i-1,j)\Big) \\
& + c(s_1^+ + s_2^+) - c(s_1^- + s_2^+) - c(s_1^+ + s_2^-) + c(s_1^- + s_2^-) \\
& \geq 0,
\end{aligned}
$$

where the terms 1-5 are non-negative by property 2, term 6 is non-negative by property 2 of inductive hypothesis and term 7 is non-negative since $c(\cdot)$ is convex and $s_2^+ \geq s_2^-$ by inductive hypothesis and Lemma 3.3.8. For the case when $i = 1$ set $s_1^- = 0$ and if $j = 0$ set $s_2^- = 0$, in all such cases the above inequality still holds.

We next show the proof for property 3(a), the proof for property 3(b) is similar. First let $i \geq 2$ and suppose s^+, s^- be the optimal resource allocation decision for $k+1$ stage problem when the state is $(i+1, j)$ and $(i-1, j+1)$ respectively. If $s_2^+ \geq s_2^-$, then using the potentially sub-optimal allocation (s_1^-, s_2^-) when the state is (i, j). The FHOE (3.8) yield,

$$\begin{aligned}
\Delta^1 v_{\alpha,k+1}(i+1, j) &= (v_{\alpha,k+1}(i+1, j) - v_{\alpha,k+1}(i, j) \\
&\geq h_1 + \lambda_1 \Delta^1 v_{\alpha,k}(i+2, j) + \lambda_2 \Delta^1 v_{\alpha,k}(i+1, j+1) + c(s_1^+ + s_2^+) - c(s_1^- + s_2^-) \\
&+ (A_1\mu_1 - s_1^+\mu_1) v_{\alpha,k}(i+1, j) + s_1^+\mu_1 v_{\alpha,k}(i, j) \\
&+ (A_2\mu_2 - s_2^+\mu_2) v_{\alpha,k}(i+1, j) + s_2^+\mu_2 v_{\alpha,k}(i+1, j-1) \\
&- (A_1\mu_1 - s_1^-\mu_1) v_{\alpha,k}(i, j) - s_1^-\mu_1 v_{\alpha,k}(i-1, j) \\
&- (A_2\mu_2 - s_2^+\mu_2) v_{\alpha,k}(i, j) - s_2^-\mu_2 v_{\alpha,k}(i, j-1)
\end{aligned}$$

Similarly using the potentially sub-optimal decision (s_1^+, s_2^+) when the state is $(i, j+1)$ yields,

$$\begin{aligned}
\Delta^1 v_{\alpha,k+1}(i, j+1) &= (v_{\alpha,k+1}(i, j+1) - v_{\alpha,k+1}(i-1, j+1) \\
&\geq h_1 + \lambda_1 \Delta^1 v_{\alpha,k}(i+1, j+1) + \lambda_2 \Delta^1 v_{\alpha,k}(i, j+2) + c(s_1^+ + s_2^+) - c(s_1^- + s_2^-) \\
&+ (A_1\mu_1 - s_1^+\mu_1) v_{\alpha,k}(i, j+1) + s_1^+\mu_1 v_{\alpha,k}(i-1, j+1) \\
&+ (A_2\mu_2 - s_2^+\mu_2) v_{\alpha,k}(i, j+1) + s_2^+\mu_2 v_{\alpha,k}(i, j) \\
&- (A_1\mu_1 - s_1^-\mu_1) v_{\alpha,k}(i-1, j+1) - s_1^-\mu_1 v_{\alpha,k}(i-2, j+1) \\
&- (A_2\mu_2 - s_2^+\mu_2) v_{\alpha,k}(i-1, j+1) - s_2^-\mu_2 v_{\alpha,k}(i-1, j)
\end{aligned}$$

Combining the two inequalities we have,

$$\begin{aligned}
\Delta^1 v_{\alpha,k+1}(i+1,j) - \Delta^1 v_{\alpha,k+1}(i,j+1) \geq\ & \lambda_1\Big(\Delta^1 v_{\alpha,k}(i+2,j) - \Delta^1 v_{\alpha,k}(i+1,j+1)\Big) \\
& + \lambda_2\Big(\Delta^1 v_{\alpha,k}(i+1,j+1) - \Delta^1 v_{\alpha,k}(i,j+2)\Big) \\
& + (A_1\mu_1 - s_1^+\mu_1)\Big(\Delta^1 v_{\alpha,k}(i+1,j) - \Delta^1 v_{\alpha,k}(i,j+1)\Big) \\
& + s_1^-\mu_1\Big(\Delta^1 v_{\alpha,k}(i,j) - \Delta^1 v_{\alpha,k}(i-1,j+1)\Big) \\
& + (A_2\mu_2 - s_2^+\mu_2)\Big(\Delta^1 v_{\alpha,k}(i+1,j) - \Delta^1 v_{\alpha,k}(i,j+1)\Big) \\
& + s_2^+\mu_2\Big(\Delta^1 v_{\alpha,k}(i+1,j-1) - \Delta^1 v_{\alpha,k}(i,j)\Big) \\
& + (s_2^+ - s_2^-)\mu_2\Big(\Delta^2 v_{\alpha,k}(i-1,j+1) - \Delta^2 v_{\alpha,k}(i,j)\Big) \\
& \geq 0,
\end{aligned}$$

where the 1-5 are non-negative by property 3a) of the inductive hypothesis, term 6 is non-negative by property 2 of inductive hypothesis and term 7 is non-negative by property 3b) of the inductive hypothesis and since we assumed that $s_2^+ \geq s_2^-$. For the case when $i = 1$, set $s_1^- = 0$ and if $j = 0$ set $s_2^- = 0$, in all such cases the above inequality still holds. If $s_2^- \geq s_2^+$ then using s_2^- at station 2 in state (i, j) and s_2^+ at station 2 in state $(i, j+1)$ leads to a similar proof. Result follows since $v_\alpha^k \uparrow v_\alpha$. ■

Proof of Lemma 3.3.14: For property 1, we use induction. Let $v_{\alpha,0}(i, j) = 0$ for all $i \geq 0, j \geq 0$, the result holds trivially for $k = 0$. For the inductive step, suppose $v_{\alpha,k}(x)$ satisfies the result for all $i \geq 0, j \geq 1$. Suppose a^* is the optimal decision for the $k+1$ stage problem when the state is $(i+1, j-1)$. First assume that $i, j \geq 1$, the FHOE (3.15) yields,

$$\begin{aligned}(\alpha+1)v_{\alpha,k+1}(i+1,j-1) &= (i+1)h_1 + (j-1)h_2 + \lambda_1 v_{\alpha,k}(i+2,j-1) + \lambda_2 v_{\alpha,k}(i+1,j) \\ &\quad + c(a^*) + \mu a^* v_{\alpha,k}(i,j-1) + (S\bar{\mu} - \mu a^*)v_{\alpha,k}(i+1,j-1) \\ &\geq ih_1 + jh_2 + \lambda_1 v_{\alpha,k}(i+1,j) + \lambda_2 v_{\alpha,k}(i,j+1) \\ &\quad + c(a^*) + \mu a^* v_{\alpha,k}(i-1,j) + (S\bar{\mu} - \mu a^*)v_{\alpha,k}(i,j) \\ &\geq (\alpha+1)v_{\alpha,k}(i,j) \end{aligned} \tag{B.1}$$

Where, the first inequality follows from the induction hypothesis and the assumption that $h_1 \geq h_2$. The second inequality follows since a^* is potentially sub-optimal for the state (i, j). If $i = 0$, then it is optimal to serve queue 2 in state (i, j), so we can replace the term corresponding to service completion in the second inequality $v_{\alpha,k}(i-1, j)$ by $v_{\alpha,k}(i, j-1)$. Property 2 now follows since $v_{\alpha,k+1}(x)) \uparrow v_\alpha(x)$ for all $x \in \mathbb{X}$.

Property 2 of the discounted cost value function follows directly from Statement (2) of Theorem 3.3.10 and the form of the dynamic programming mapping (3.31).

For the proof of Property 3 (convexity), note that since convexity follows from Properties 4 and 5, it suffices to show the last two properties. We again use induction and the FHOE (3.15) to prove properties 4 and 5. Let $v_{\alpha,0}(i, j) = 0$ for all $i \geq 0, j \geq 0$, the result holds trivially for $k = 0$. For the inductive step, suppose $v_{\alpha,k}(i, j)$ satisfies Properties 3,4 and 5. We start with property 4. Suppose $i \geq 2$ and let a^1, a^0 be the optimal resource allocation decision for the $k+1$ stage problem when the initial state is $(i, j+1)$ and $(i-1, j)$ respectively. Suppose we use the potentially sub-optimal allocation a^0 when the state is $(i-1, j+1)$. The

FHOE (3.15) yield,

$$
\begin{aligned}
(\alpha+1)\Delta^1 v_{\alpha,k+1}(i,j+1) &= (\alpha+1)(v_{\alpha,k+1}(i,j+1) - v_{\alpha,k+1}(i-1,j+1)) \\
&\geq h_1 + \lambda_1 \Delta^1 v_{\alpha,k}(i+1,j+1) + \lambda_2 \Delta^1 v_{\alpha,k}(i,j+2) + (1-\lambda_1-\lambda_2)\Delta^1 v_{\alpha,k}(i,j+1) \\
&\quad + c(a^1) - a^1\mu\Delta^1 v_{\alpha,k}(i,j+1) - c(a^0) + a^0\mu\Delta^1 v_{\alpha,k}(i-1,j+1)
\end{aligned}
$$

Similarly using the potentially sub-optimal decision a^1 when the state is (i, j) yields,

$$
\begin{aligned}
(\alpha+1)\Delta^1 v_{\alpha,k+1}(i,j) &= (\alpha+1)(v_{\alpha,k+1}(i,j) - v_{\alpha,k+1}(i-1,j) \\
&\leq h_1 + \lambda_1 \Delta^1 v_{\alpha,k}(i+1,j) + \lambda_2 \Delta^1 v_{\alpha,k}(i,j+1) + (1-\lambda_1-\lambda_2)\Delta^1 v_{\alpha,k}(i,j) \\
&\quad + c(a^1) - a^1\mu\Delta^1 v_{\alpha,k}(i,j) - c(a^0) + a^0\mu\Delta^1 v_{\alpha,k}(i-1,j)
\end{aligned}
$$

Combining the two inequalities we have,

$$
\begin{aligned}
(\alpha+1)(\Delta^1 v_{\alpha,k+1}(i,j+1) - \Delta^1 v_{\alpha,k+1}(i,j)) &\geq \lambda_1\Big(\Delta^1 v_{\alpha,k}(i+1,j+1) - \Delta^1 v_{\alpha,k}(i+1,j)\Big) \\
&\quad + \lambda_2\Big(\Delta^1 v_{\alpha,k}(i,j+2) - \Delta^1 v_{\alpha,k}(i,j+1)\Big) \\
&\quad + (1-\lambda_1-\lambda_2-a^1\mu)\Big(\Delta^1 v_{\alpha,k}(i,j+1) - \Delta^1 v_{\alpha,k}(i,j)\Big) \\
&\quad + a^0\mu\Big(\Delta^1 v_{\alpha,k}(i-1,j+1) - \Delta^1 v_{\alpha,k}(i-1,j)\Big) \\
&\geq 0,
\end{aligned}
$$

where the last inequality follows by Property 4 of the inductive hypothesis. Now suppose $i = 1$, using the same approach as before we have,

$$\begin{aligned}(\alpha+1)(\Delta^1 v_{\alpha,k+1}(1, j+1) - \Delta^1 v_{\alpha,k+1}(1, j)) &\geq \lambda_1\Big(\Delta^1 v_{\alpha,k}(2, j+1) - \Delta^1 v_{\alpha,k}(2, j)\Big) \\ &\quad + \lambda_2\Big(\Delta^1 v_{\alpha,k}(1, j+2) - \Delta^1 v_{\alpha,k}(1, j+1)\Big) \\ &\quad + (1 - \lambda_1 - \lambda_2 - a^1\mu)\Big(\Delta^1 v_{\alpha,k}(1, j+1) - \Delta^1 v_{\alpha,k}(1, j)\Big) \\ &\quad + a^0\mu\Big(\Delta^2 v_{\alpha,k}(0, j+1) - \Delta^2 v_{\alpha,k}(0, j)\Big) \\ &\geq 0,\end{aligned}$$

where the second inequality follows by Properties 3 and 4 of the inductive hypothesis.

We next show the proof for Property 5(b), the proof for Property 5(a) is similar. Let $i \geq 1$ and suppose a^1, a^0 be the optimal resource allocation decision for $k+1$ stage problem when the state is $(i, j+1)$ and $(i+1, j-1)$ respectively. If $a^0 \geq a^1$, then using the potentially sub-optimal allocation a^1 when the state is (i, j). The FHOE (3.15) yield,

$$\begin{aligned}(\alpha+1)\Delta^2 v_{\alpha,k+1}(i, j+1) &= (\alpha+1)(v_{\alpha,k+1}(i, j+1) - v_{\alpha,k+1}(i, j) \\ &\geq h_2 + \lambda_1\Delta^2 v_{\alpha,k}(i+1, j+1) + \lambda_2\Delta^2 v_{\alpha,k}(i, j+2) + (S\bar{\mu} - a^1\mu)v_{\alpha,k}(i, j+1) \\ &\quad + a^1\mu v_{\alpha,k}(i-1, j+1) - (S\bar{\mu} - a^1\mu)v_{\alpha,k}(i, j) - a^1\mu v_{\alpha,k}(i-1, j)\end{aligned}$$

Similarly using the potentially sub-optimal decision a^0 when the state is $(i+1, j)$ yields,

$$\begin{aligned}(\alpha+1)\Delta^2 v_{\alpha,k+1}(i+1, j) &= (\alpha+1)(v_{\alpha,k+1}(i+1, j) - v_{\alpha,k+1}(i+1, j-1) \\ &\leq h_2 + \lambda_1\Delta^2 v_{\alpha,k}(i+2, j) + \lambda_2\Delta^2 v_{\alpha,k}(i+1, j+1) + (S\bar{\mu} - a^0\mu)v_{\alpha,k}(i+1, j) \\ &\quad + a^0\mu v_{\alpha,k}(i, j) - (S\bar{\mu} - a^0\mu)v_{\alpha,k}(i+1, j-1) - a^0\mu v_{\alpha,k}(i, j-1)\end{aligned}$$

Combining the two inequalities we have,

$$\begin{aligned}(\alpha+1)(\Delta^2 v_{\alpha,k+1}(i,j+1)-\Delta^2 v_{\alpha,k+1}(i+1,j)) \geq\ & \lambda_1\Big(\Delta^2 v_{\alpha,k}(i+1,j+1)-\Delta^2 v_{\alpha,k}(i+2,j)\Big)\\ &+\lambda_2\Big(\Delta^2 v_{\alpha,k}(i,j+2)-\Delta^2 v_{\alpha,k}(i+1,j+1)\Big)\\ &+(S\bar{\mu}-a^1\mu)\Big(\Delta^2 v_{\alpha,k}(i,j+1)-\Delta^2 v_{\alpha,k}(i+1,j)\Big)\\ &+(a^0-a^1)\mu\Big(\Delta^2 v_{\alpha,k}(i+1,j)-\Delta^2 v_{\alpha,k}(i,j)\Big)\\ &+a^1\mu\Big(\Delta^2 v_{\alpha,k}(i-1,j+1)-\Delta^2 v_{\alpha,k}(i,j)\Big)\\ \geq\ & 0,\end{aligned}$$

where the last inequality follows because $a^0 \geq a^1$ and by Properties 4 and 5(b) of the inductive hypothesis. Now suppose $i = 0$, using the same approach as before we have,

$$\begin{aligned}(\alpha+1)(\Delta^2 v_{\alpha,k+1}(0,j+1)-\Delta^2 v_{\alpha,k+1}(1,j)) \geq\ & \lambda_1\Big(\Delta^2 v_{\alpha,k}(1,j+1)-\Delta^2 v_{\alpha,k}(2,j)\Big)\\ &+\lambda_2\Big(\Delta^2 v_{\alpha,k}(0,j+2)-\Delta^2 v_{\alpha,k}(1,j+1)\Big)\\ &+(S\bar{\mu}-a^1\mu)\Big(\Delta^2 v_{\alpha,k}(0,j+1)-\Delta^2 v_{\alpha,k}(1,j)\Big)\\ &+(a^0-a^1)\mu\Big(\Delta^2 v_{\alpha,k}(1,j)-\Delta^2 v_{\alpha,k}(0,j)\Big)\\ \geq\ & 0,\end{aligned}$$

where the second inequality follows because $a^0 \geq a^1$ and by properties 4 and 5(b) of inductive hypothesis. If $a^1 \geq a^0$, then using the potentially sub-optimal decision a^0 in state (i, j) and the decision a^1 in state $(i + 1, j)$ yields,

$$\begin{aligned}(\alpha+1)(\Delta^2 v_{\alpha,k+1}(i,j+1)-\Delta^2 v_{\alpha,k+1}(i+1,j)) \geq\ & \lambda_1\Big(\Delta^2 v_{\alpha,k}(i+1,j+1)-\Delta^2 v_{\alpha,k}(i+2,j)\Big)\\ &+\lambda_2\Big(\Delta^2 v_{\alpha,k}(i,j+2)-\Delta^2 v_{\alpha,k}(i+1,j+1)\Big)\\ &+(S\bar{\mu}-a^1\mu)\Big(\Delta^2 v_{\alpha,k}(i,j+1)-\Delta^2 v_{\alpha,k}(i+1,j)\Big)\\ &+(a^1-a^0)\Big(\Delta^1 v_{\alpha,k}(i+1,j)-\Delta^1 v_{\alpha,k}(i,j)\Big)\\ &+a^1\Big(\Delta^2 v_{\alpha,k}(i-1,j+1)-\Delta^2 v_{\alpha,k}(i,j)\Big)\\ \geq\ & 0,\end{aligned}$$

where the second inequality follows because $a^1 \geq a^0$ and by Properties 3(a) and 5(b) of inductive hypothesis. For the case when $i = 0$, a similar approach yields,

$$\begin{aligned}
(\alpha+1)(\Delta^2 v_{\alpha,k+1}(0, j+1) - \Delta^2 v_{\alpha,k+1}(1, j)) &\geq \lambda_1\Big(\Delta^2 v_{\alpha,k}(1, j+1) - \Delta^2 v_{\alpha,k}(2, j)\Big) \\
&\quad + \lambda_2\Big(\Delta^2 v_{\alpha,k}(0, j+2) - \Delta^2 v_{\alpha,k}(1, j+1)\Big) \\
&\quad + (S\bar{\mu} - a^1\mu)\Big(\Delta^2 v_{\alpha,k}(0, j+1) - \Delta^2 v_{\alpha,k}(1, j)\Big) \\
&\quad + (a^1 - a^0)\mu\Big(v_{\alpha,k}(1, j-1) - v_{\alpha,k}(0, j)\Big) \\
&\geq 0,
\end{aligned}$$

where the second inequality follows because $a^1 \geq a^0$ and by Properties 1 and 3(b) of the inductive hypothesis.

■

BIBLIOGRAPHY

[1] Amazon ec2. `http://aws.amazon.com/ec2/`.

[2] Google appengine. `https://cloud.google.com/appengine/`.

[3] Oracle: On demand scaling. `http://www.oracle.com/technetwork/topics/cloud/paas-on-demand-scale-098391.html`.

[4] Hyun-Soo Ahn and Mark E Lewis. Flexible Server Allocation and Customer Routing Policies for Two Parallel Queues When Service Rates Are Not Additive. *Operations Research*, 61(2):344–358, April 2013.

[5] E. Altman. *Constrained Markov Decision Process*. CRC Press, 1999.

[6] B. Ata. Dynamic Power Control in a Wireless Static Channel Subject to a Quality-of-Service Constraint. *Operations Research*, 53(5):842–851, September 2005.

[7] B. Ata and K. E. Zachariadis. Dynamic power control in a fading downlink channel subject to an energy constraint. *Queueing Systems*, 55(1):41–69, December 2006.

[8] J. S. Baras, A. J. Dorsey, and A. M. Makowski. Two Competing Queues with Linear Costs and Geometric Service Requirements: The c-u Rule Is Often Optimal. *Advances in Applied Probability*, 17(1):186, March 1985.

[9] Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Zomaya. A Taxonomy and Survey of Energy- Efficient Data Centers and Cloud Computing Systems. *Engineering*, (December 2010):1–24, 2011.

[10] R.A. Berry. Power and delay trade-offs in fading channels. PhD Thesis, 2000.

[11] D.P. Bertsekas. *Dynamic Programming and Optimal Control-Vol II*. Athena Scientific, 1995.

[12] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[13] C. Buyukkoc, P. Varaiya, and Jean Walrand. The c-u rule revisited. *Advances in Applied Probability*, 17:237–238, 1985.

[14] T.B. Crabill. Optimal Control of a Maintenance System with Variable Service Rates. *Operations Research*, 22(4):736–745, July 1974.

[15] J.G. Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *The Annals of Applied Probability*, 40(11):49–77, 1995.

[16] J.G. Dai and S.P. Meyn. Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Transactions on Automatic Control*, 40(11):1889–1904, 1995.

[17] Hyun-soo Ahn David L. Kaufman and Mark E. Lewis. On the introduction of an agile, temporary workforce into a tandem queueing system. *Queueing Systems: Theory and Applications*, 51(1-2):135–171, 2005.

[18] Izak Duenyas and Mark P. Van Oyen. Stochastic scheduling of parallel queues with set-up costs. *Queueing Systems*, 19(4):421–444, December 1995.

[19] Izak Duenyas and Mark P. Van Oyen. Heuristic Scheduling of Parallel Heterogeneous Queues with Set-Ups. *Management Science*, 42(6):814–829, June 1996.

[20] E.A. Feinberg and M.E. Lewis. Optimality Inequalities for Average Cost Markov Decision Processes and the Stochastic Cash Balance Problem. *Mathematics of Operations Research*, 32(4):769–783, November 2007.

[21] V.S. Frost and B. Melamed. Modeling For Telecommunications Networks. *IEEE Communications Magazine*, (March):70–81, 1994.

[22] Anshul Gandhi, Mor Harchol-Balter, and Michael a. Kozuch. Are sleep states effective in data centers? *2012 International Green Computing Conference, IGCC 2012*, 2012.

[23] J.M. George and J.M. Harrison. Dynamic Control of a Queue with Adjustable Service Rate. *Operations Research*, 49(5):720–731, September 2001.

[24] L.V. Green and P.J. Kolesar. On the accuracy of the simple peak hour approximation for Markovian queues. *Management science*, 41(8):1353–1370, 1995.

[25] Donald Gross, John F. Shortle, James M. Thompson, and Carl M. Harris. *Fundamentals of Queueing Theory*. Wiley-Interscience, New York, NY, USA, 4th edition, 2008.

[26] V. Gupta, M. Harchol-Balter, A. Scheller Wolf, and U. Yechiali. Fundamental characteristics of queues with fluctuating load. *ACM SIGMETRICS Performance Evaluation Review*, 34(1):203, June 2006.

[27] J. Michael Harrison. A Priority Queue with Discounted Linear Costs. *Operations Research*, 23(2):260–269, April 1975.

[28] J. Michael Harrison. Dynamic Scheduling of a Multiclass Queue: Discount Optimality. *Operations Research*, 23(2):270–282, April 1975.

[29] H. Heffes and D. Lucantoni. A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *Selected Areas in Communications, IEEE Journal on*, 4(6):856–868, 1986.

[30] Micha Hofri and Keith W Ross. On the Optimal Control of Two Queues with Server Setup Times and Its Analysis. *SIAM Journal on Computing*, 16(2):399–420, April 1987.

[31] D.L. Kaufman and M.E. Lewis. Machine maintenance with workload considerations. *Naval Research Logistics*, 54(7):750–766, October 2007.

[32] J. Keilson and A. Kester. Monotone matrices and monotone Markov processes 1. *Stochastic Processes and their Applications*, 5(3):231–241, 1977.

[33] G. Koole. Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, 1(x):1–15, 1998.

[34] Ger Koole. Assigning a single server to inhomogeneous queues with switching costs. *Theoretical Computer Science*, 182(1-2):203–216, August 1997.

[35] Chiunlin Lim and Ao Tang. Dynamic speed scaling and load balancing of interconnected queues. In *Information Theory and Applications Workshop (ITA), 2011*, pages 1–10, Feb 2011.

[36] S.A. Lippman. Applying a new device in the optimization of exponential queuing systems. *Operations Research*, 23(4):687–710, 1975.

[37] Zhen Liu, Philippe Nain, and Don Towsley. On optimal polling policies. *Queueing Systems*, 11(1-2):59–83, 1992.

[38] W.A. Massey and W. Whitt. Peak congestion in multi-server service systems with slowly varying arrival rates. *Queueing Systems*, 25(1):157–172, 1997.

[39] Peter Mell and Timothy Grance. The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology. *Nist Special Publication*, 145:7, 2011.

[40] L. Muscariello, M. Meillia, M. Meo, M.A. Marsan, and R.L. Cigno. An MMPP-based hierarchical model of Internet traffic. *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, pages 2143–2147 Vol.4, 2004.

[41] Philippe Nain and Don Towsley. Optimal scheduling in a machine with stochastic varying processing rate. *IEEE Transctions On Automatic Control*, 39(9):1853–1855, 1994.

[42] Philippe Nain, Pantelis Tsoucas, and Jean Walrand. Interchange Arguments in Stochastic Scheduling. *Journal of Applied Probability*, 26(4):815, December 1989.

[43] Pantelis Tsoucas Philippe Nain and Jean Walrand. Interchange arguments in stochastic scheduling. *Journal of Applied Probability*, 26:815–826, 1989.

[44] W.B. Powell. *Approximate Dynamic Programming Solving the curses of dimensionality*. Wiley; 2 edition, 2011.

[45] N.U. Prabhu and Y. Zhu. Markov-modulated queueing systems. *Queueing Systems*, 5(1-3):215–245, November 1989.

[46] M. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2011.

[47] Martin I. Reiman and Lawrence M. Wein. Dynamic Scheduling of a Two-Class Queue with Setups. *Operations Research*, 46(4):532–547, August 1998.

[48] L.I. Sennott. Average Cost Semi-Markov Decision Processes and the Control of Queueing Systems. *Probability in the Engineering and Informational Sciences*, 3(02):247, July 1989.

[49] L.I. Sennott. *Stochastic Dynamic Programming and the Control of Queueing Systems (Wiley Series in Probability and Statistics)*. John Wiley and Sons, Inc., New York, 1999.

[50] S. Shah-Heydari. MMPP modeling of aggregated ATM traffic. *Conference Proceedings. IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No.98TH8341)*, 1:129–132, 1998.

[51] Mitrani I. Slegers J. and Thomas N. Evaluating the optimal server allocation policy for clusters with on/off sources. *Performance Evaluation*, 66(8):453–467, August 2009.

[52] S. Stidham Jr and R.R. Weber. Monotonic and insensitive optimal policies for control of queues with undiscounted costs. *Operations Research*, 37(4):611–625, 1989.

[53] E. Uysal-Biyikoglu, B. Prabhakar, and a. El Gamal. Energy-efficient packet transmission over a wireless link. *Networking, IEEE/ACM Transactions on*, 10(4):487–499, 2002.

[54] Josh Whitney and Pierre Delforge. Data Center Efficiency Assessment Scaling Up Energy Efficiency Across the Data Center Industry : Evaluating Key Drivers and Barriers. Technical Report August, NRDC, 2014.

[55] A. Wierman, L. L. H. Andrew, and A. Tang. Power-Aware Speed Scaling in Processor Sharing Systems. In *IEEE INFOCOM 2009 - The 28th Conference on Computer Communications*, number Section IV, pages 2007–2015. IEEE, April 2009.

[56] Ran Yang, Sandjai Bhulai, and Rob van der Mei. Optimal resource allocation for multiqueue systems with a shared server pool. *Queueing Systems*, 68(2):133–163, 2011.

[57] U. Yechiali. A queuing-type birth-and-death process defined on a continuous-time Markov chain. *Operations Research*, 21(2):604–609, 1973.

[58] S. Yoon and M.E. Lewis. Optimal Pricing and Admission Control in a Queueing System with Periodically Varying Parameters. *Queueing Systems*, 47(3):177–199, July 2004.