Metrics, Fundamental Trade-offs and Control Policies for Delay-sensitive Applications in Volatile Environments

by

Ali ParandehGheibi

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

Doctor of Philosophy
in
Electrical Engineering and Computer Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2012

Author .....
Department of Electrical Engineering and Computer Science
January 10, 2012

Certified by .....
Muriel Médard
Professor
Thesis Supervisor

Certified by .....
Asuman Ozdaglar
Associate Professor
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Students

Metrics, Fundamental Trade-offs and Control Policies for Delay-sensitive Applications in Volatile Environments

by

Ali ParandehGheibi

S.M., Electrical Engineering and Computer Science, Massachusetts Institute of Technology (2008)

B.Sc., Electrical Engineering, Sharif University of Technology (2006)

Submitted to the Department of Electrical Engineering and Computer Science
on January 10, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

With the explosion of consumer demand, media streaming will soon be the dominant type of Internet traffic. Since such applications are intrinsically delay-sensitive, the conventional network control policies and coding algorithms may not be appropriate tools for data dissemination over networks. The major issue with design and analysis of delay-sensitive applications is the notion of *delay*, which significantly varies across different applications and time scales.

We present a framework for studying the problem of media streaming in an unreliable environment. The focus of this work is on end-user experience for such applications. First, we take an analytical approach to study fundamental rate-delay-reliability trade-offs in the context of media streaming for a single receiver system. We consider the probability of interruption in media playback (buffer underflow) as well as the number of initially buffered packets (initial waiting time) as the Quality of user Experience (QoE) metrics. We characterize the optimal trade-off between these metrics as a function of system parameters such as the packet arrival rate and the file size, for different channel models. For a memoryless channel, we model the receiver's queue dynamics as an M/D/1 queue. Then, we show that for arrival rates slightly larger than the play rate, the minimum initial buffering required to achieve certain level of interruption probability remains bounded as the file size grows. For the case where the arrival rate and the play rate match, the minimum initial buffer size should scale as the square root of the file size. We also study media streaming over channels with memory, modeled using Markovian arrival processes. We characterize the optimal trade-off curves for the infinite file size case, in such Markovian environments.

Second, we generalize the results to the case of multiple servers or peers streaming to a single receiver. Random linear network coding allows us to simplify the packet

selection strategies and alleviate issues such as duplicate packet reception. We show that the multi-server streaming problem over a memoryless channel can be transformed into a single-server streaming problem, for which we have characterized QoE trade-offs.

Third, we study the design of media streaming applications in the presence of multiple heterogeneous wireless access methods with different access costs. Our objective is to analytically characterize the trade-off between usage cost and QoE metrics. We model each access network as a server that provides packets to the user according to a Poisson process with a certain rate and cost. User must make a decision on how many packets to buffer before playback, and which networks to access during the playback. We design, analyze and compare several control policies. In particular, we show that a simple Markov policy with a threshold structure performs the best. We formulate the problem of finding the optimal control policy as a Markov Decision Process (MDP) with a probabilistic constraint. We present the Hamilton-Jacobi-Bellman (HJB) equation for this problem by expanding the state space, and exploit it as a verification method for optimality of the proposed control policy.

We use the tools and techniques developed for media streaming applications in the context of power supply networks. We study the value of storage in securing reliability of a system with uncertain supply and demand, and supply friction. We assume storage, when available, can be used to compensate, fully or partially, for the surge in demand or loss of supply. We formulate the problem of optimal utilization of storage with the objective of maximizing system reliability as minimization of the expected discounted cost of blackouts over an infinite horizon. We show that when the stage cost is linear in the size of the blackout, the optimal policy is myopic in the sense that all shocks are compensated by storage up to the available level of storage. However, when the stage cost is strictly convex, it may be optimal to curtail some of the demand and allow a small current blackout in the interest of maintaining a higher level of reserve to avoid a large blackout in the future. Finally, we examine the value of storage capacity in improving system's reliability, as well as the effects of the associated optimal policies under different stage costs on the probability distribution of blackouts.

Thesis Supervisor: Muriel Médard
Title: Professor

Thesis Supervisor: Asuman Ozdaglar
Title: Associate Professor

To my wife Shabnam,
my parents Saeed and Zahra,
and my sister Marzieh.

Acknowledgments

I have been extremely fortunate that my life path has crossed some of the most brilliant people in the world. The people whose support, good heartedness and wisdom have played an important role in completion of this thesis.

First of all I would like to express my gratitude to my advisors Prof. Muriel Médard and Prof. Asu Ozdaglar for all of their selfless help and support. I am extremely grateful for them believing in me and taking a chance on me ever since, or even before, I started my graduate studies at MIT. I have been very fortunate to receive advice from both Asu and Muriel, each teaching me different technical knowledge and insight, research styles, and teaching practices. I particularly appreciate their efforts for providing me the proper environment and freedom to grow as an independent researcher, and their encouragement for trying new ideas, while subtly guiding me through uncertain roads of discovery. Without all their guidance and help, I could not have accomplished any of the works throughout my graduate studies.

I also would like to acknowledge Asu and Muriel's dedication in reading line by line and draft after draft of my papers. However, Ph.D studies is not all about publishing papers. It is mostly about development of one's character as a critical thinker, self-driven and independent researcher. On this road, I have been very fortunate to have Muriel and Asu as my mentors and role-models, both in academic and personal life.

I am especially grateful for Asu and Muriel's patience and support that expand well beyond their responsibilities as research advisors. Their positive thinking and encouragements in difficult times gave me the confidence to carry on.

I would like to thank Prof. Stephen Boyd and Prof. Srikant for serving on my dissertation committee. I am very grateful for the time they selflessly spent on reading and improving the content of this thesis. With their deep insight and useful comments, they helped to broaden my vision in wider perspective.

I have had the pleasure of working with Prof. Daron Acemoglu of MIT Economics department. Daron is not only one of the greatest minds of our time, and a true visionary, he is also extremely dedicated in working closely with students. His hard

work and insightful advice has been a great inspiration for me.

I have had the great fortune of interacting with several other professors during my graduate studies. I would like to thank Prof. Nancy Lynch for teaching me a different perspective in my research studies, and for her useful suggestions and meticulous comments. I would also like to thank Prof. Srinivas Shakkottai who provided several useful suggestions for different parts of this work. Finally, I would like to thank Dr. Mardavij Roozbehani, who with his broad knowledge, has helped me at different stages of my Ph.D. works. The last part of this manuscript would not be complete without his help. I am deeply grateful to the faculty LIDS, RLE and CSAIL, especially Devavrat Shah, John Tsitsiklis, Munther Dahleh, Dimitri Bertsekas, Lizhong Zheng, Sash Megretski and Dina Katabi for offering amazing courses, helping me with all of my random questions and creating such a great learning environment at MIT.

I am deeply indebted to my wife Shabnam Ghadarghadr. This thesis would not be complete without her help and support. Her continuous care, understanding and sacrifice helped me tackle obstacles of life with great confidence. I would like to thank Shabnam for her patience, trust, passion and unconditional love. Not only has she helped me through her tolerance and encouragement, but also she has helped me with the technical challenges of my work, when I was completely hopeless. She is the one who believes in my dreams. With her support, I feel I can accomplish anything in my life.

I had the opportunity of being a member of two amazing groups, who always treated me like family. I am very grateful for all my seniors for mentoring me through the graduate school process. I would like to thank all of my friends, officemates and colleagues in LIDS, RLE and CSAIL, with whom I had many fruitful discussions and collaborations. I have certainly enjoyed their company. In particular, I would like to thank MinJi Kim, Atilla Eryilmaz, Jay Kumar Sundararajan, Ozan Candogan, Kostas Bimpikis, Leonardo Urbina, Soheil Feizi, Daniel Lucani, Michael Kilian, Shirley Shi, Ilan Lobel, Amir Ali Ahmadi, Parikshit Shah, Mesrob Ohannessian, Majid Khabbazian, Kerim Fouli, Ulric Ferner, Alireza Tahbaz, Noah Stein, Ermin Wei, Kimon Drakopoulos, Arman Rezaee, Flaávio du Pin Calmon, Wifei Zeng, Jason

Cloud, Tong Wang, Marie-Jose Montpetit, Danail Traskov, Abbasali Makhdoumi, Fang Zhao, Georgios Angelopoulos, Surat Teerapittayanon, Nadia Fawaz, Desmond Lun, Chris Ng, Ishai Menache, Paul Njoroge, Ercan Yildiz, Vitaly Abdrashitov, Azarakhsh Malekian, Annie Chen, Elie Adam, Emmanuel Abbe and Sertac Karaman. A very special thanks to Michael Lewy, Lynne Dell, Jennifer Donovan, Janet Fischer, Brian Jones, Laura Dargus, David Foss, Sukru Cinar and all other administrative staff, who have been a tremendous help all these years. Aside from my friends and colleagues at the institute, I would like to thank all my friends at Sharif University who are now spread all around the world. They helped create the most amazing experience for me during every single year of my undergraduate and graduate life.

Last, but not least, I would like to express my heartfelt gratitude to my parents, Saeed and Zahra, my sister, Marzieh, mother-in-law, Jila, brother-in-law, Hamed, sister-in-laws, Nastaran and Yasi for all of their love and support. I would like to thank my parents, who have been my first teachers and have always supported me with their sacrifice, love, patience and encouragement. It has been really hard for them to be apart from their child for many years, but they embraced it with patience, just for the sake of my success. Without their help and selfless support, I did not stand a chance of overcoming obstacles of life.

The work in this thesis was partially supported by subcontract # 18870740-37362-C issued by Stanford University and supported by the Defense Advanced Research Projects Agency (DARPA), and by the National Science Foundation (NSF) under Grant No. CNS-0627021.

Contents

List of Figures

Chapter 1

Introduction

Media streaming is fast becoming the dominant application on the Internet [1]. The predictions [2] show that by 2014, the various forms of video (TV, VoD, Internet Video, and P2P) will exceed 91 percent of global consumer traffic. Figure 1-1 illustrates the growth of IP traffic for different applications based on Cisco's forecast. The popularity of such media transfers has been accompanied by the growing usage of wireless handheld devices as the preferred means of media access. It is expected that such media streaming would happen in both a device to device (D2D) as well as in a base-station to device fashion, and both the hardware and applications needed for such communication schemes are already making an appearance [3, 4].

Media streaming applications are intrinsically delay-sensitive. Hence, they need to be managed differently from the traditional less delay-sensitive applications such as Web, Email and file downloads. Most of the current approaches for providing a reasonable Quality of Service (QoS) for streaming applications are based on resource over-provisioning. The basic philosophy behind this approach is that higher bandwidth and more reliable communication links result in smaller delays and better QoS. Therefore, users need a broadband connection or a dedicated reliable channel to have a reasonable Quality of Experience (QoE). Nevertheless, with rapid growth of IP traffic and resource limitations specially in the wireless domain, this approach seems to be less plausible over time.

In this work, we address this problem by focusing specifically on end-user per-

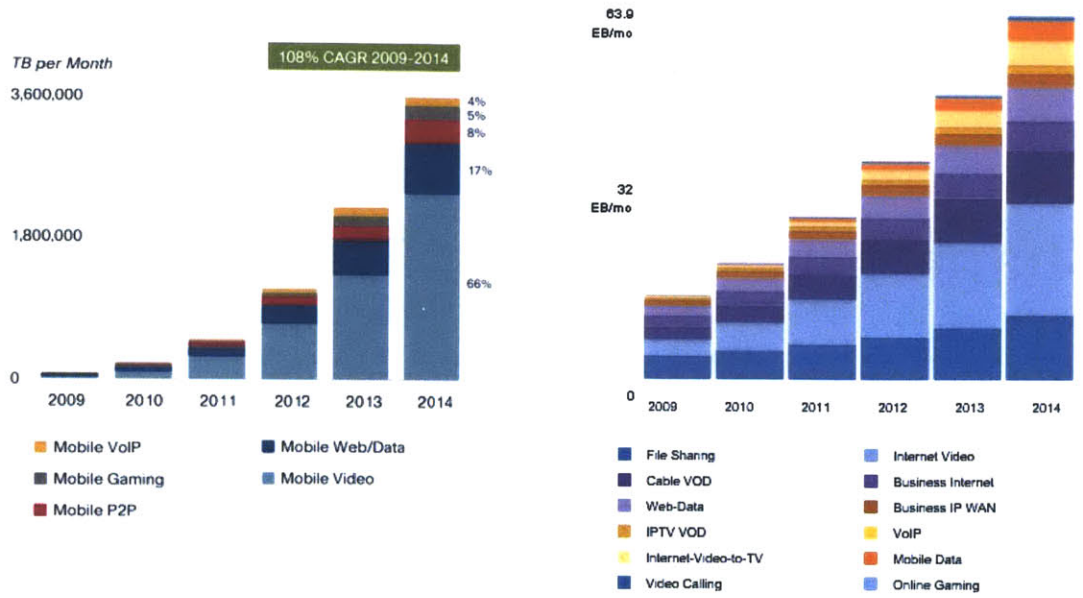

Figure 1-1: Cisco Visual Networking Index (June 2010): By 2014 video will be 91 percent of all consumer IP traffic and 66 percent of mobile data traffic.

ceptions of delay-sensitive applications. We pay special attention to communication and control techniques that are specifically designed with Quality of user Experience (QoE) in mind. The goal, on one hand, is to make the optimal use of the limited and possibly unreliable resources to provide a seamless experience for the user. On the other hand, we would like to provide a tool for the service providers to improve their service delivery (specifically for delay-sensitive applications) in the most economical way. Our contributions are summarized in the following.

1.1 End-user Quality of Experience Metrics

Media streaming is generally accomplished by dividing the media file into *blocks,* which are then further divided into packets for transmission. After each block is received, it can be played out by the receiver. In order to ensure smooth sequential playback, a fresh block must be received before the current block has been played. If such a fresh block is not available the playback stops, causing a negative user experience.

For most of media streaming applications e.g. Internet video, TV, Video on

Demand (VoD), the user may tolerate some initial buffering delay *before* the media playback, so that he or she has a seamless experience *throughout* the playback. Hence, there is a trade-off between the initial waiting time and likelihood of playback interruptions. Two of the key Quality of user Experience (QoE) metrics that we consider in this work, are the initial waiting time and the probability of having any interruptions throughout the media playback. Here, the waiting time captures the *delay* aspect of the user experience, and the interruption probability captures the *reliability* aspect of the experience.

We would like to emphasize that these QoE metrics are significantly different from those that consider per packet deadlines as the delay, and the number of missed deadlines as the reliability metric. These metrics are more appropriate for highly interactive applications such as video conferencing. Even though video communications traffic is experiencing a fast growth, these applications still form a small fraction of the total IP traffic [2]. The metrics concerning per packet delays are too restrictive for most of the media streaming applications, because they do not take into account the slack obtained by the initial buffering. Therefore, they are generally more difficult to satisfy unless by over-provisioning. The QoE metrics that we introduce are more realistic and less restrictive, but still strong enough to guarantee a reasonable user experience. For example, if the probability of playback interruption is guaranteed to be less than one percent, it means out of 100 media files streamed to the end-user, approximately 99 of them will be played with no interruptions.

We may consider other less restrictive QoE metrics such as the number or duration of interruptions (stalls) in media playback. For example, the user may tolerate a few interruptions in media playback as long as they are not too long or too frequent. Other QoE metrics of interest are the resolution (clarity) of the media playback as well as temporal variations of these features.

1.2 Fundamental QoE Trade-offs for Single-server Streaming Systems

Once we identify the appropriate QoE metrics, our first goal is to understand the fundamental trade-offs among end-user rate, delay and reliability metrics in the context of media streaming. For communication over a noisy channel (physical layer), the following equation captures the essence of such trade-offs from an information theoretic point of view:

$$\text{Probability of block decoding error} = e^{-E(R)\cdot(\text{block length})}, \tag{1.1}$$

where $E(R)$ is the error exponent (reliability function), which depends on rate R and the properties of the channel. The block length is considered as the *delay* metric, and probability of recovering the original block of symbols captures the *reliability* of communication. The delay-reliability trade-off is governed by the error exponent, which is an increasing function of the gap between the *rate*, R, and the channel's capacity. In this work, we establish similar rate-delay-reliability trade-offs for media streaming applications.

The main objective of this part is to characterize the amount of buffering needed for a target probability of playback interruption over the duration of the playback. If the packets arrive at the receiver deterministically, the required initial buffering is zero when the packet arrival rate is larger than the playback rate, and it grows linearly with the file size when the arrival rate is smaller than the playback rate. However, since most of the communication links of interest are noisy and unreliable, packets cannot be obtained deterministically. Thus, our question is *how much should we buffer prior to playback in order to account for channel variations?*

We first consider the problem of streaming a media file of finite size from a single server to a single receiver. In order to take into account the channel variations such as packet erasures, we model the packet arrival process at the receiver using a Poisson process of rate R. Since the media file is generally played in a deterministic fashion,

we model the receiver's buffer as an M/D/1 queue. We then provide upper and lower bounds on the minimum initial buffering required so that the playback interruption probability is below a desired level. The optimal trade-off between the initial buffering and the interruption probability depends on the file size as well as the playback rate compared to the arrival rate of the packets. Our bounds are asymptotically tight as the file size tends to infinity. Moreover, when the arrival rate and the play rate match, we show that the minimum initial buffer size grows as the square-root of the file size. In this case, the amount of buffering is solely for alleviating the stochastic variations of the communication channel, so that the end-user will have a seamless experience. If the arrival rate is slightly larger than the play rate, the minimum initial buffering for a given interruption probability remains bounded as the file size grows. In particular, for the infinite file size case, we establish the following relation

$$\text{Probability of interruption} = e^{-I(R)\cdot(\text{initial buffering})}, \tag{1.2}$$

where we define $I(R)$ as the *interruption exponent* or reliability function in analogy with (1.1). In (1.2), the reliability metric (interruption probability) is related to the delay metric (initial buffering) via the reliability function $I(R)$.

The fundamental relation in (1.2) is not restricted to the case of Poisson arrival processes. We may establish such result for a more general class of channels. In this work, we also *explicitly* characterize the interruption exponent for Poisson arrival models as well as channels with memory, which are modeled using Markov modulated arrival processes. For channels with memory, we show that the behavior of the interruption exponent is not only governed by the average packet arrival rate, but also by the mixing rate of the underlying Markov chain.

Explicit characterization of the interruption exponents is also extremely valuable from system designer's point of view. This allows the designer to observe how sensitive end-user experience is to changes in allocated communication resources. Therefore, having an insight on the level of customer satisfaction, the system operator can decide whether provisioning more resources is economically reasonable. We shall discuss this

aspect in the context of streaming from multiple servers.

1.3 Multi-server Streaming Systems and Network Coding

We address the problem of streaming delay-sensitive information from multiple servers to a single receiver (user). Each sever can be a wireless access point or another peer operating as a server. We consider a model in which the communication link between the receiver and each server is unreliable, and hence, it takes a random period of time for each packet to arrive at the receiver from the time that the packet is requested from a particular server. One of the major difficulties with such multi-server systems is the packet tracking and duplicate packet reception problem, i.e., the receiver need to keep track of the index of the packets it is requesting from each server to avoid requesting duplicate packets. Since the requested information are delay sensitive, if a requested packet does not arrive within some time interval, the receiver need to request the packet from another server. This may eventually result in receiving duplicate packets and waste of the resources. We address this issue and show that using random linear network coding (RLNC) across packets within each block of the media file we can alleviate this issue. This technique allows that, with high probability, no redundant information will be delivered to the receiver.

Another important issue regarding multi-server, in particular peer-to-peer (P2P) systems, is the packet selection (piece selection) strategy. This issue arises in situations where some of the servers do not carry all of the pieces (packets) of a media file, hence, they may not be able to help the receiver with its missing piece. The piece selection strategy is to decide which packet to request from a particular server (peer) among all possible packets that are available at that server, so that to avoid wasting the communication resources on contacting a server that may not have any *useful* pieces. In this work, we use random linear network coding to address this problem. Here, the receiver instead of requesting a particular packet, asks for a generic packet

within a particular block, and the server transmits a random linear combination of the available packets in that block. This greatly simplifies the piece selection strategy and the need for tracking the packets, while effectively guaranteeing delivery of useful packets from each server with high probability.

We would like to emphasize that one of the critical roles of network coding techniques in this work, other than improving efficiency, is their simplicity and ability to greatly simplify the communication models, so that we can focus on end-user metrics and trade-offs. For example, if each server can effectively transmit packets according to an independent Poisson process, using RLNC we can merge these processes into one Poisson process of sum rate. Hence, the system model boils down to a single-server system for which the optimal single-server QoE trade-offs results apply.

1.4 Control Policies for Streaming in Heterogeneous Multi-server Systems

When there are multiple networks that can be used to access a particular piece of content (e.g. from a base station or a peer device) each device must take decisions on associating with one or more such access networks. However, the costs of different access methods might be different. For example, accessing the base station of a cellular network can result in additional charges per packet, while it might be possible to receive the same packets from the access point of a Wireless Local Area Network (WLAN) or another device with a lower cost or possibly for free (see Figure 1-2). Further, the cost of communication might be mitigated by the initial amount of buffering before playback. This adds another dimension to the problem that end-user is facing. That is certain levels of user satisfaction can only be achieved by paying a premium for extra resource availability. Figure 1-3 illustrates a conceptual three-dimensional cost-delay-reliability trade-off curve.

The objective of this part is to understand the trade-off between initial waiting time, and the usage cost for attaining a target probability of interruption. We consider

Figure 1-2: Media streaming from heterogeneous servers.

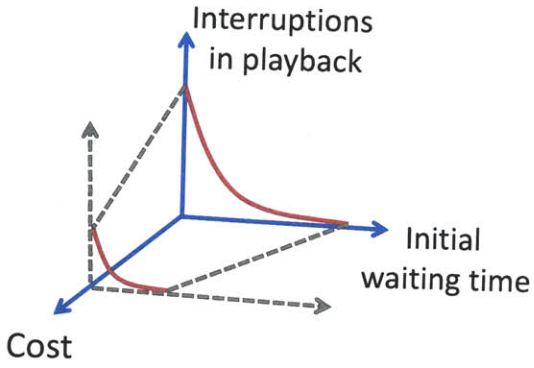

Figure 1-3: Trade-off between the achievable QoE metrics and cost of communication.

a system wherein network coding is used to ensure that packet identities can be ignored, and packets may potentially be obtained from two class of servers with different rates of transmission. The wireless channel is unreliable, and we assume that each server can deliver packets according to a Poisson process with a known rate. Further, the costs of accessing the two servers are different; for simplicity we assume that one of the servers is free. Thus, *our goal is to develop a control policy that switches between the free and the costly servers in order to attain the target QoE metrics at the lowest cost.*

We study several classes of server selection policies. Using the QoE trade-offs for a single-server system, we obtain a lower bound on the cost of offline policies that do not observe the trajectory of packets received. We show that such policies have a threshold form in terms of the time of association with the costly server. Using the offline algorithm as a starting point, we develop an *online* algorithm with lower cost that has a threshold form – both free and costly servers are used until the queue length reaches a threshold, followed by only free server usage. We then develop an online algorithm in which the risk of interruption is spread out across the trajectory. Here, only the free server is used whenever the queue length is above a certain threshold, while both servers are used when the queue length is below the threshold. The threshold is designed as a function of the initial buffer size and the

desired interruption probability.

We formulate the problem of finding the optimal network association policy as a Markov Decision Process (MDP) with a probabilistic constraint. Similar to the Bellman equation proposed by Chen [5] for a discrete-time MDP with probabilistic constraints, we write the Hamilton-Jacobi-Bellman (HJB) equation for our continuous-time problem by proper state space expansion. Using a guess and check approach, we obtain a solution for the HJB equation for the fluid approximation model. We characterize the optimal policy and show that the it is Markov with respect to the queue-length process at the receiver. In particular, we establish the optimality of the online threshold policy with a properly designed threshold value.

1.5 Analytical Tools and Techniques

Even though the main application of our work is media streaming, the tools that we have developed to analyze our system from this new perspective are not limited to such applications. In order to characterize the fundamental trade-offs between QoE metrics, especially for the case of finite file size, we need to perform a transient analysis of the system. The set of techniques that we developed to provide tight bounds on interruption probabilities are different from the traditional approaches in the Queueing Theory literature. The main focus of that body of work is to understand system's stability conditions as well as steady state behavior, e.g. queue length, waiting time, etc. These steady state metrics cannot be easily transformed to reasonable QoE metrics for streaming applications because of their delay-sensitive nature.

The techniques developed in this work can be employed for analysis of other delay-sensitive applications such as cloud-based services and applications. Moreover, our methods could be useful in other areas such as insurance and financial industries where statistical methods are used for risk assessment. The questions of interest in these industries include computing the likelihood of an insurance or financial institution facing bankruptcy, and similar ruin related measures.

Another set of tools that we used in this work include dynamic programming

approach to Markov decision problems with probabilistic constraints. Since the probabilistic constraint is over the whole horizon of the problem, we need to expand the state space to decouple the state evolution and cost. This technique can be of particular interest for Inventory Control problems. Such problems include decision making by a firm on how much to order in any time period to meet the customer's demand. It is generally undesirable to have a situation where the inventory is empty due to high customer demand or late arrival of the ordered product. This situation is similar to an interruption in media playback. Instead of the conventional approach of back-ordering and assigning an artificial penalty for back-ordered products, we can use the probabilistic constraints in our formulation to avoid stock-outs. The dynamic programming approach used in this work allows us to develop inventory control policies of the threshold form, which are close to optimal and simple enough from practical point of view.

1.6 Application to Power Systems in Volatile Environments

We employ the analytical tools developed motivated by media streaming applications in the context of power systems in volatile environments. Supply and demand in electric power networks are subject to exogenous, impulsive, and unpredictable shocks due to generator outages, failure of transmission equipments or unexpected changes in weather conditions. On the other hand, environmental causes along with price pressure have led to a global trend in large-scale integration of renewable resources with stochastic output. This is likely to increase the magnitude and frequency of impulsive shocks to the supply side of the network. The questions that we are interested are: what is the value of energy storage in mitigating volatility of supply and demand, and what are the fundamental limits that cannot be overcome by storage due to physical ramp constraints for charging the storage, and finally, what are the impacts of different control policies on system reliability, for instance, on the expected

cost or the probability of large blackouts?

We focus on the reliability value of storage, defined as the maximal improvement in system reliability as a function of storage capacity. Two metrics for quantifying reliability in a system are considered: The first is the expected long-term discounted cost of blackouts (cost of blackouts (COB) metric), and the second is the probability of loss of load by a certain amount or less.

We model the system as a supply-demand model that is subject to random arrivals of energy deficit shocks, and a storage of limited capacity, with a ramp constraint on charging, but no constraint on discharging. The storage may be used to partially or completely mask the shocks to avoid blackouts. We formulate the problem of optimal storage management as the problem of minimization of the COB metric, and provide several characterizations of the optimal cost function. By ignoring other factors such as the environment, cost of energy or storage, we characterize the value of storage purely from a reliability perspective, and examine the effects of physical constraints on system reliability. Moreover, for a general convex stage cost function, we present various structural properties of the optimal policy.

In particular, we prove that for a linear stage cost, a *myopic* policy which compensates for all shocks regardless of their size by draining from storage as much as possible, is optimal. However, for nonlinear stage costs where the penalty for larger blackouts is significantly higher, the myopic policy is not optimal. Intuitively, the optimal policy is inclined to mitigate large blackouts at the cost of allowing more frequent small blackouts. Our numerical results confirm this intuition. We further investigate the value of additional storage under different control policies, and for different ranges of system parameters. Our results suggest that if the ratio of the average rate of deficit shocks to ramp constraints is sufficiently large, there is a critical level of storage capacity above which, the value of having additional capacity quickly diminishes. When this ratio is significantly large, there seems to be another critical level for storage size below which, storage capacity provides very little value. Finally, we investigate the effect of storage size and volatility of the demand/supply process on the probability of large blackouts under various policies. We observe that for all

control policies, there appears to be a critical level of storage size, above which the probability of suffering large blackouts diminishes quickly.

1.7 Related Work

The set of works related to this thesis spans several distinct areas of the literature. One of the major difficulties in the literature is the notion of *delay*, which greatly varies across different applications and time scales at which the system is modeled. The role of delay-related metrics has been extensive in the literature on Network Optimization and Control. Neely [6, 7] employs Lyapunov Optimization techniques to study the delay analysis of stochastic networks, and its trade-offs with other utility functions. Other related works such as [8], [9] take the flow-based optimization approach, also known as Network Utility Maximization (NUM), to maximize the *delay-related* utility of the users. Closer to our work is the one by Hou and Kumar [10] that considers per-packet delay constraints and successful delivery ratio as user experience metrics. Such flow-based approaches are essentially operating at the steady state of the system, and fail to capture the end-user experience for delay-sensitive applications of interest.

Media streaming, particularly in the area of P2P networks has attracted significant recent interest. For example, works such as [11, 12, 13, 14] develop analytical models on the trade-off between the steady state probability of missing a block and buffer size under different block selection policies. Unlike our model, they consider live streaming, e.g. video conferencing, with deterministic channels.

There are two main approaches to P2P video streaming, namely, (i) using push-based multicast trees, and (ii) using pull-based mesh P2P networks. Push-based multicast trees require that each entering user join one or more multicast trees [15, 16, 17]. Each block is pushed along a multicast tree to ensure that each user obtains blocks sequentially and with an acceptable delay. However, such an approach often involves excessive infrastructural overheads, and peer churn causes inefficiencies [18]. Pull-based mesh P2P has recently seen significant usage as a means of video delivery. Here peers maintain a playback buffer and pull blocks from each other. The approach

is similar to the popular BitTorrent protocol [19], which makes use of a full mesh with a subset of peers being exposed to each peer. This approach has been used in many systems such as CoolStreaming [20], PPLive [21], QQLive [22] and TVAnts [23]. A more recent modification is to use random linear network coding techniques [24] to make block selection simpler [25, 26, 27] in the wired and wireless context. RLNC allows for a push-based mechanism without requiring the overhead of a tree-based push strategy. We use the same idea to ensure that packets can be received from multiple sources without the need to coordinate the exact identities of the packets from each. However, we focus on content that is at least partially cached at multiple locations, and must be streamed over one or more unreliable channels. Further, our analysis is on transient effects—we are interested in the first time that media playback is interrupted as a function of the initial amount of buffering. Also related to our work is [28], which considers two possible wireless access methods (WiFi and UMTS) for file delivery, assuming particular throughput models for each access method. In contrast to this work, packet arrivals are stochastic in our model, and our streaming application requires hard constraints on quality of user experience.

The techniques we use to compute the optimal trade-off curves in the infinite file size case, are related to those used in the literature of Ruin Theory [29], which study insurer's vulnerability to insolvency. In particular, Reinhard [30] employs a system of integro-differential equations to characterize the non-ruin probabilities of an insurer with constant premium rates and exponentially distributed claim amounts in a Markovian environment. For the finite file size case, we need to characterize hitting probabilities of crossing a time-varying threshold for which such methods are not effective. Our work could be of independent interest since it provides novel techniques for characterizing the trade-offs with finite file sizes.

Another body of related work is the literature on constrained Markov decision processes. There are two main approaches to these problems. Altman [31], Piunovskiy [32, 33] and Feinberg and Shwartz [34] take a Convex Analytic approach leading to Linear Programs for obtaining the optimal policies. On the other hand, Chen [5], Chen and Blankenship [35], Piunovskiy [36] use the more natural and straightforward

Dynamic Programming approach to characterize all optimal policies. These works mainly focus on different variations of the discrete-time Markov decision processes. In this work, we take the dynamic programming approach for the control of a continuous-time Markovian jump process. Further, we employ Stochastic Calculus techniques used in treatment of Stochastic Control problems [37] to properly characterize the optimal control policies.

Related to energy storage and power systems, recent works have examined the effects of ramp constraints on the economic value of storage [38]. Herein, our focus is on reliability. Prior research on using queueing models for characterization of system reliability, particularly in power systems, has been reported in [39] and [40]. Similar models and concepts exist in the queueing theory literature [41], [42], perhaps with different application contexts. Despite similarities, our model is different than those of [39], [40] in many ways. We assume that the storage capacity (reserve in their model) is fixed and find the optimal policy for withdrawing from storage (consuming from reserve), as opposed to always draining the reserve and optimizing the capacity. Another difference is that our model of uncertainty is a compound poisson process instead of the brownian motion used in [39], [40]. We show that the myopic policy of always draining storage to mask every energy deficit shock is not optimal for strictly convex costs, and investigate the effects of nonlinear stage costs (strictly convex cost of blackouts) on the optimal policy and the statistics of blackouts.

Notation. *Throughout the thesis,* $\mathbb{I}_A$ *denotes the indicator function of a set* A*. We use the notation* $\mathbb{R}$ *for real numbers, and* $\mathbf{Pr}\{\cdot\}$ *and* $\mathbf{E}[\cdot]$ *for probability and expectation operators, respectively. We also denote the minimum of two scalars* x*,* y *by* $x \wedge y$*. The operator* $[x]^+ = \max\{0, x\}$ *is the projection operator onto the nonnegative orthant.*

1.8 Thesis Outline

The rest of this thesis is organized as follows. In Chapter 2, we formally define the QoE metrics and characterize the optimal QoE trade-off curves for various channel models in a point-to-point system. In Chapter 3, we generalize the results to a system

with multiple servers and discuss network coding techniques. Chapter 4 is dedicated to the problem of media streaming from heterogeneous servers. We present and compare several server association policies. The dynamic programming approach for characterization of the optimal control policy is discussed in this chapter. In Chapter 5, we present our results on the reliability value of energy storage for power systems in volatile environments. Finally, Chapter 6 presents a summary of this thesis with pointers for potential extensions in the future.

Chapter 2

Fundamental QoE Trade-offs for Media Streaming in Single-server Systems

We start by investigating the problem of streaming a media file from a single server to a single user. The goal is to have a proper model of factors that affect Quality of user Experience, and understand the fundamental trade-off between these metrics as a function of system parameters. We start with the simplest model of an unreliable communication channel and generalize the results to other channel models and multi-server systems.

2.1 System Model and QoE Metrics

Consider a single server (peer) streaming a media file to a single receiver (user) over an unreliable channel. Time is continuous, and we assume it takes a random amount of time for the server to successfully deliver each packet to the receiver. For this part, we assume that the packet arrival process at the receiver is a Poisson process of rate R. This corresponds to the case of a memoryless channel. We shall study channels with memory in the subsequent parts. We assume that the media file consists of F packets. We normalize the playback rate to one, i.e., it takes one unit of time to play

a single packet. We also assume that the parameter R is known at the receiver, which first buffers D packets from the beginning of the file, and then starts the playback. The dynamics of the receiver's buffer size[1] $Q(t)$ can be described as follows

$$Q(t) = D + A(t) - t, \tag{2.1}$$

where D is the initial buffer size and $A(t)$ is a Poisson process of rate R. We declare an interruption in playback when the buffer size reaches zero *before* completely downloading the media file. More precisely, let

$$\tau_e = \inf\{t : Q(t) \leq 0\}, \qquad \tau_f = \inf\{t : Q(t) \geq F - t\}, \tag{2.2}$$

where τ_f corresponds to time of completing the file download, because we have already played τ_f packets and the buffer contains the remaining $F - \tau_f$ packets to be played. The media streaming is interrupted if and only if $\tau_e < \tau_f$.

We consider the following metrics to quantify Quality of user Experience (QoE). The first metric is the initial waiting time before the playback starts. This is captured by the initial buffer size D. The second metric is the probability of interruption during the playback denoted by

$$p(D) = \mathbf{Pr}\{\tau_e < \tau_f\}, \tag{2.3}$$

where τ_e and τ_f are defined in (2.2). In our model, the user expects to have an interruption-free experience with probability higher than a desired level $1 - \epsilon$. Note that there is a fundamental trade-off between the interruption probability ϵ and the initial buffer size D. For example, owing to the randomness of the arrival process, in order to have zero probability of interruption, it is necessary to fully download the file, i.e., $D = F$. Nevertheless, we need to buffer only a small fraction of the file if user tolerates a positive probability of interruption. These trade-offs and their relation to system parameters R and F are addressed in the following.

[1]We use the terms "buffer size" and "queue length" interchangeably throughout this work.

2.2 Optimal QoE Trade-offs

We would like to obtain the smallest initial buffer size so that the interruption probability is below a desired level ϵ, which is denoted by

$$D^*(\epsilon) = \min\{D \geq 0 : p(D) \leq \epsilon\}, \tag{2.4}$$

where $p(D)$ is the interruption probability defined in (2.3). Note that in general $p(D)$ and hence $D^*(\epsilon)$ depend on the arrival rate R and the file size F which are assumed to to be known. In the following we characterize the optimal trade-off between the initial buffer size and the interruption probability by providing bounds on $D^*(\epsilon)$. An upper bound (achievability theorem) on $D^*(\epsilon)$ is particularly useful, since it provides a sufficient condition for desirable user experience. A lower bound (converse theorem) of $D^*(\epsilon)$ is helpful to show that the provided upper bound is close to the exact value. The proofs of the main theorems are included in the appendix of this chapter.

Theorem 2.2.1. [Achievability] *Let $D^*(\epsilon)$ be defined as in (2.4), and $I(R)$ be the largest root of $\gamma(r) = r + R(e^{-r} - 1)$, i.e.,*

$$I(R) = \sup\{r : \gamma(r) = 0\}. \tag{2.5}$$

(a) *For all $R > 1$,*

$$D^*(\epsilon) \leq \frac{1}{I(R)} \log\left(\frac{1}{\epsilon}\right). \tag{2.6}$$

(b) *For all $0 \leq R \leq 1 + \left(\frac{1}{2F}\log\left(\frac{1}{\epsilon}\right)\right)^{\frac{1}{2}}$,*

$$D^*(\epsilon) \leq \min\left\{F(1-R) + \left(2FR\log\left(\frac{1}{\epsilon}\right)\right)^{\frac{1}{2}}, \frac{1}{I(R)}\log\left(\frac{1}{\epsilon}\right)\right\}. \tag{2.7}$$

When the arrival rate R is smaller than one (the playback rate), the upper bound in Theorem 2.2.1 consists of two components. The first term, $F(1-R)$, compensates the expected number of packets that are required by the end of $[0, F]$ period. The second component, $\left(2FR\log\left(\frac{1}{\epsilon}\right)\right)^{\frac{1}{2}}$, compensates the randomness of the arrivals to

avoid interruptions with high probability. Note that this term increases by decreasing the maximum allowed interruption probability, and it would be zero for a deterministic arrival process. For the case when the arrival rate is larger than the playback rate, the minimum required buffer size does not grow with the file size. This is so since the buffer size in (3.1) has a positive drift. Hence, if there is no interruption at the beginning of the playback period, it becomes more unlikely to happen later.

In the following, we show that the upper bounds presented in Theorem 2.2.1 are *asymptotically tight*, by providing lower bounds on the minimum required buffer size $D^*(\epsilon)$, for different regimes of the arrival rate R. Let us first define the notion of a tight bound.

Definition 2.2.1. Let $\hat{D}$ be a lower or upper bound of the $D^*(\epsilon)$, which depends on the file size F. The bound $\hat{D}$ is an *asymptotically tight* bound if $\frac{|\hat{D}-D^*(\epsilon)|}{D^*(\epsilon)}$ vanishes as F goes to infinity.

Theorem 2.2.2. [Converse] *Let $D^*(\epsilon)$ be defined as in (2.4), and $I(R)$ be given by (2.5). Then*

(a) *For all $R > 1$,*

$$D^*(\epsilon) \geq -\frac{1}{I(R)} \log\left(\epsilon + 2e^{-\frac{(R-1)^2}{4(R+1)}F}\right). \tag{2.8}$$

(b) *For each $R \leq 1$ and $\epsilon \leq \frac{1}{16}$, if $F \geq C \log\left(\frac{1}{\epsilon}\right)$ for a constant C, then*

$$D^*(\epsilon) \;\geq\; F(1-R) + \frac{1}{2}\left(2FR\log\left(\frac{1}{\epsilon}\right)\right)^{\frac{1}{2}}. \tag{2.9}$$

Note that the inequality (2.9) in part (b) of Theorem 2.2.2 does not hold for all ϵ. In fact, we can show that $D^*(\epsilon) < F(1-R)$ for a large interruption probability ϵ. In the extreme case $\epsilon = 1$, it is clear that $D^*(\epsilon) = 0$. Nevertheless, since we are interested in *avoiding* interruptions, we do not study this regime of the interruption probabilities. Comparing the lower bounds obtained in Theorem 2.2.2 with the upper bounds obtained in Theorem 2.2.1, we observe that they demonstrate a similar behavior as a function of the parameters F and R.

Corollary 2.2.1. *The upper bounds and lower bounds of $D^*(\epsilon)$ given by Theorems 2.2.1 and 2.2.2 are asymptotically tight, if $R > 1$, or $R < 1$ and $\epsilon \leq \frac{1}{16}$.*

Corollary 2.2.2. *Let $p(D)$ be the interruption probability defined in (2.3). For the case of $R > 1$, and infinite file size, $F = \infty$, we have*

$$p(D) = \mathbf{Pr}\Big\{ \min_{t \geq 0} Q(t) \leq 0 \Big\} = e^{-I(R)D}, \qquad \text{for all } D \geq 0, \tag{2.10}$$

where $I(R)$ is defined in (2.5).

Proof. The proof simply follows from Theorems 2.2.1 and 2.2.2, and continuity of the probability measure. □

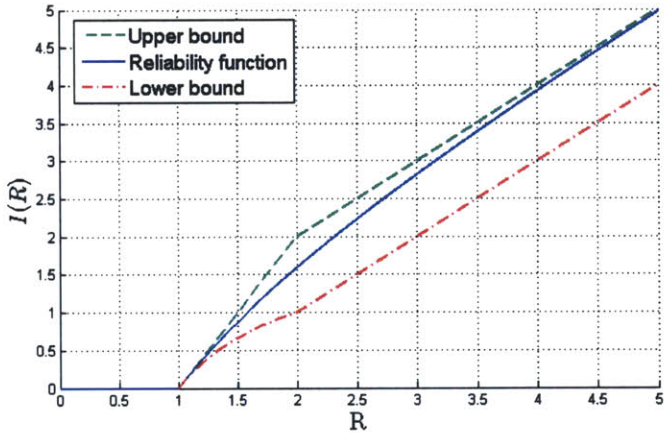

Figure 2-1: The reliability function (interruption exponent) defined in (2.5) for the Poisson arrival process. Simple lower and upper bounds are given by Lemma 2.6.1.

Corollary 2.2.2 relates the *reliability* of media playback (interruption probability) to *delay* in media playback (initial buffer size) via the *reliability function*, $I(R)$. The reliability function or *interruption exponent* depends on the *rate* and distribution of the arrival process, i.e., properties of the communication channel. Figure 2-1 plots the reliability function $I(R)$ as a function of the arrival rate. Observe that as the ratio of the arrival rate and the playback rate increases, so does the reliability function. When arrival and playback rates match, the reliability function is zero. This behavior is reminiscent of the error exponent of a noisy communication channel, i.e.,

the error exponent is zero when the rate is equal to the capacity, and is increasing in the distance between the communication rate and the capacity. In the context of media streaming, Corollary 2.2.2 provides an analogue of the error-exponent characterizations that capture the delay-rate-reliability trade-off of block channel codes used for communication over noisy channels.

Thus far, we have studied the QoE trade-offs for media streaming applications with Poisson arrivals. This model does not capture the burstiness of the traffic often associated with correlated losses over a wireless channel. In the following, we generalize the results of this section to the case where packets arrive according to a Markovian process. These results are analogous to error exponent characterization of channels with memory such as the Gilbert-Elliot channel [43].

2.3 QoE Trade-offs for Bursty Traffic (Markovian Channels)

Consider a media streaming scenario from a server to a single receiver. In this part, we focus on the case where the media file size is infinite, and the packet arrival rate is strictly larger than the playback rate (normalized to one). The packet arrival process is governed by a two-state continuous time Markov process depicted in Figure 2-2. For each state i, λ_i denotes the transition rate from state i to the other state. We may verify that the stationary distribution of this Markov process is given by $\pi_i = 1 - \frac{\lambda_i}{\lambda_1+\lambda_2}$, $i = 1, 2$.

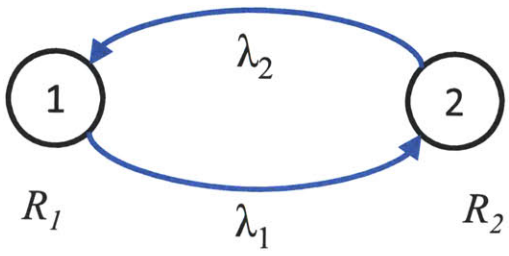

Figure 2-2: Two-state Markov process used to model the burstiness of packet arrivals. λ_1 and λ_2 denote transition rates.

Let the packets arrive at the receiver according to some stationary stochastic process with rate R_i, when the underlying Markov process is in state i. Throughout this work, we assume that R_1 and R_2 are such that the average arrival rate is larger than the playback rate, i.e.,

$$\bar{R} = \sum_{i=1,2} \pi_i R_i = \frac{\lambda_2 R_1 + \lambda_1 R_2}{\lambda_1 + \lambda_2} > 1. \tag{2.11}$$

For $\bar{R} < 1$, the receiver's queue-length has a negative drift and interruption occurs almost surely for any finite initial buffer size. For a Markovian channel, the interruption probability not only depends on D, but also on the initial state. Let $p_i(D)$ be the interruption probability, when the initial state of the underlying Markov process is i, and the initial buffer size is D. In the following, we study QoE trade-offs for various scenarios of Markovian packet arrival processes.

2.3.1 Markovian Channels with Deterministic Arrivals

In this part, we focus on the case where the packet arrival process is deterministic with rate R_i, when the underlying two-state Markov process is in state i. The trade-off between the interruption probability and the initial buffer size is characterized next.

Theorem 2.3.1. *Consider the Markov process depicted in Figure 2-2. Let the arrival process at the receiver be deterministic with rate R_i, when in state i. If $R_1 > 1 > R_2$ and (2.11) holds, then*

$$p_1(D) = \frac{\rho_1}{\rho_2} e^{-(\rho_2 - \rho_1)D}, \quad p_2(D) = e^{-(\rho_2 - \rho_1)D}, \tag{2.12}$$

where $\rho_i = \frac{\lambda_i}{|R_i - 1|}$, for $i = 1, 2$.

Proof. Since we are considering the infinite file size case, the interruption probability starting from any particular time, only depends on the queue-length and the state of

the Markov process at that time. For any $D, h \geq 0$, write

$$p_i(D) = \lambda_i h p_{-i}\big(D + (R_i - 1)h\big) + (1 - \lambda_i h) p_i\big(D + (R_i - 1)h\big) + o(h), \quad i = 1, 2,$$

where $-i$ denotes the index of the state other than i, and $o(h)$ denotes the terms that vanish faster than h, as h goes to zero. Dividing by h and letting h tend to 0, we get

$$\frac{\partial p_i(D)}{\partial D} = \rho_i\big(p_1(D) - p_2(D)\big), \quad i = 1, 2, \tag{2.13}$$

with the boundary condition $p_i(\infty) = 0$. The assumption in (2.11) is necessary for this boundary condition to hold. Using (2.13) together with the boundary condition, we have $p_1(D) = \frac{\rho_1}{\rho_2} p_2(D)$. Replacing this relation back in (2.13) for $i = 2$, we obtain a differential equation for $p_2(D)$

$$\frac{\partial p_2(D)}{\partial D} = \rho_2 \left(\frac{\rho_1}{\rho_2} - 1 \right) p_2(D),$$

which has the solution $p_2(D) = c_1 e^{-(\rho_2 - \rho_1)D} + c_2$. Observe that $c_2 = 0$ by the boundary condition. Since $R_2 < 1$, we also have $p_2(0) = 1$. Hence, $c_1 = 1$, that proves the claim in (2.12). □

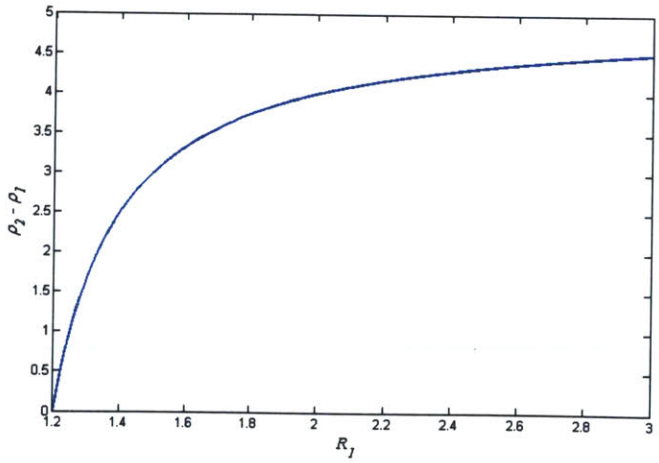

Figure 2-3: The reliability function, $\rho_2 - \rho_1$, given by Theorem 2.3.1 for the Gilbert-Elliot channel with deterministic arrivals. Here, $\lambda_1 = \lambda_2 = 1$, and $R_2 = 0.8$.

We observe similar exponential decay in the interruption probability as in Corollary 2.2.2. The interruption exponent (reliability function) given by $\rho_2 - \rho_1$ is increasing in R_1 and R_2. Figure 2-3 plots the reliability function, $\rho_2 - \rho_1$, as a function of R_1, when $R_2 = 0.8$ and $\lambda_1 = \lambda_2 = 1$. Observe that the reliability function goes to zero as the average arrival rate $\bar{R}$ approaches the playback rate. It is worth mentioning that for a fixed average arrival rate, as λ_1 and λ_2 increase the mixing time of the Markov process decreases and the hence, the arrival process tend to look deterministic with rate $\bar{R} > 1$. Therefore, larger reliability function is achieved. This behavior is similar to that of the conventional Gilbert-Elliot communication channel [43].

Note that the characterization of the trade-off curve is trivial for the case where $R_1, R_2 > 1$. In this case, $p_i(D) = 0$, for $i = 1, 2$.

2.3.2 Markovian Channels with Poisson Arrivals

We consider a two-state Markov modulated Poisson process as the packet arrival process. The Poisson arrivals allow us to model the channel variations in small time scales, while the underlying Markov process models the large scale changes in the environment. Next, we characterize the interruption probabilities as a function of the initial state and buffer size.

Theorem 2.3.2. *Consider the two-state Markov process depicted in Figure 2-2. Let the arrival process at the receiver be Poisson with rate R_i, when in state i. If the average arrival rate is larger than the playback rate, i.e., (2.11) holds, then*

$$p_1(D) = c_{11}e^{-s_1 D} + c_{12}e^{-s_2 D}, \quad p_2(D) = c_{21}e^{-s_1 D} + c_{22}e^{-s_2 D}, \tag{2.14}$$

where $s_1 < s_2$ are the positive roots of the characteristic function

$$\Phi(s) = (s + R_1(e^{-s} - 1) - \lambda_1)(s + R_2(e^{-s} - 1) - \lambda_2) - \lambda_1\lambda_2, \tag{2.15}$$

and

$$c_{i1} = \frac{s_2 + R_i(e^{-s_2} - 1)}{(s_2 - s_1) + R_i(e^{-s_2} - e^{-s_1})}, \quad c_{i2} = -\frac{s_1 + R_i(e^{-s_1} - 1)}{(s_2 - s_1) + R_i(e^{-s_2} - e^{-s_1})}. \tag{2.16}$$

Proof. We sketch the proof owing to space limitation. Similarly to the proof of Theorem 2.3.1, and [30], we obtain the following system of delay differential equations

$$\frac{\partial p_i(D)}{\partial D} = R_i\Big(p_i(D+1) - p_i(D)\Big) + \lambda_i\Big(p_{-i}(D) - p_i(D)\Big), \quad i = 1, 2. \tag{2.17}$$

Using the assumption in (2.11), we get the boundary conditions $p_1(\infty) = p_2(\infty) = 0$. Moreover, since the packets depart from the receiver's queue deterministically, we obtain the additional boundary conditions $p_1(0) = p_2(0) = 1$. We may solve the above system of differential equations using the roots of characteristic equation given by (2.15). Observe that the first set of boundary conditions imply that we only need to consider the terms that vanish as D grows to infinity. We may verify that $\Phi(s)$ has two roots $s_1, s_2 > 0$ if the condition in (2.11) holds (see Figure 2-4). Therefore, the interruption probabilities take the form of (2.14). Further, the second set of boundary conditions imply that $c_{i1} + c_{i2} = 1$ for $i = 1, 2$. We may solve for the coefficients in (2.16) by plugging (2.14) in (2.17), and using these boundary conditions. □

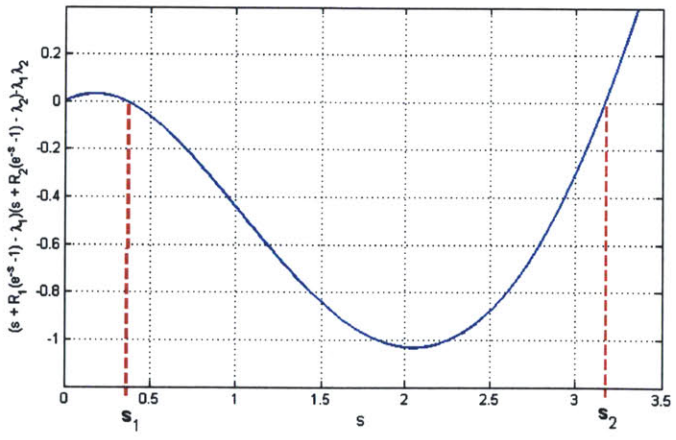

Figure 2-4: The characteristic function $\Phi(s)$ given by (2.15). s_1 denotes the dominant root.

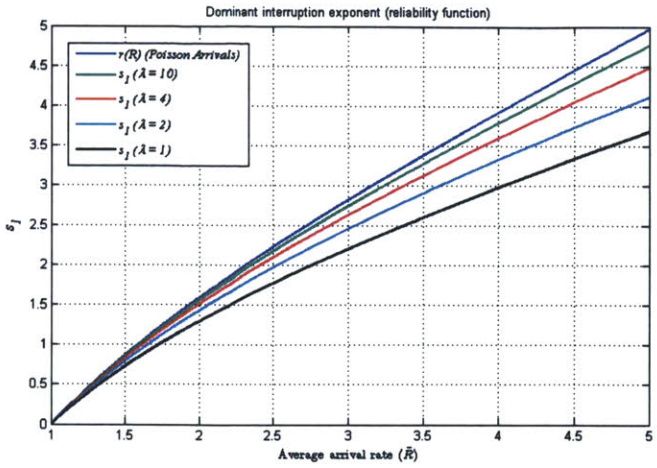

Figure 2-5: The dominant root of the characteristic function given by (2.15), as a function of the average arrival rate for different $\lambda_1 = \lambda_2 = \lambda$. Here, we select $R_1 = 1.4\bar{R}$, and $R_2 = 0.6\bar{R}$.

The behavior of the interruption probabilities given by Theorem 2.3.2 is governed by the dominant mode s_1. Even though, we cannot compute the roots of the characteristic function analytically, we may examine approximate solutions of the characteristic equation when the Markov process is fast mixing, i.e., $\lambda_1, \lambda_2 \gg 1$. Using the notation in (2.11), we can rewrite (2.15) as

$$\Phi(s) = \big(s + R_1(e^{-s} - 1)\big)\big(s + R_2(e^{-s} - 1)\big) - (\lambda_1 + \lambda_2)\big(s + \bar{R}(e^{-s} - 1)\big). \quad (2.18)$$

Now, let λ_1, λ_2 tend to infinity, while keeping R_1 and R_2 fixed. There are two possibilities for any root of $\Phi(s)$: First, the root does not grow to infinity; second, the root scales to infinity as λ_1, λ_2 grow. In the former case, we need to have the coefficient of $(\lambda_1 + \lambda_2)$ in (2.18) go to zero, i.e., for λ_1, λ_2 large enough we have $s + \bar{R}(e^{-s} - 1) \approx 0$. Therefore, using the notation in (2.5), $s = I(\bar{R})$ is an approximate root of $\Phi(s)$. Figure 2-5 plots s_1, the exact root of $\Phi(s)$, as a function of the average arrival rate, $\bar{R}$, for different mixing times. Observe that as λ grows, s_1 approaches $I(\bar{R})$, which is the reliability function for the Poisson arrival case (see Corollary 2.2.2). In the second case, we may approximate the characteristic function in (2.18) as $\Phi(s) \approx s^2 - (\lambda_1 + \lambda_2)s$. This results in $s_2 \approx \lambda_1 + \lambda_2$ that grows to infinity with λ_1

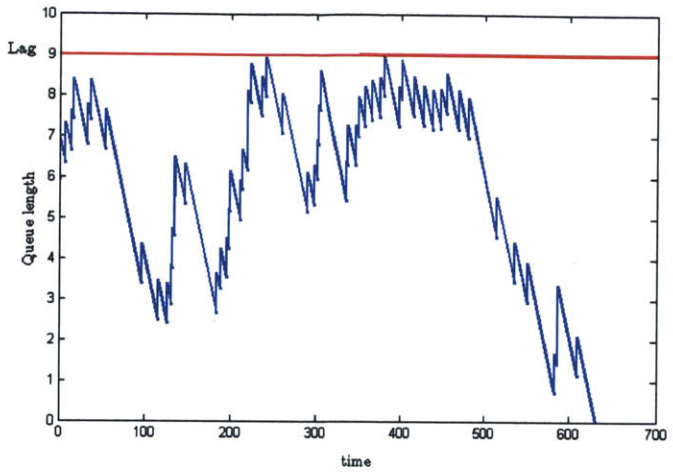

Figure 2-6: The queue dynamics for a live streaming scenario.

and λ_2. This root is significantly larger than the dominant root, s_1, which remains bounded. Finally, note that if $R_1 = R_2 = R$, we have $c_{12} = c_{22} = 0$ for all λ_1 and λ_2. This confirms the earlier result of Corollary 2.2.2.

2.4 Live Streaming Applications

Thus far, we have considered streaming scenarios where the content is completely available at the servers. Such scenarios are of particular interest for applications such as VoD (Video on Demand) and time-shifted programming (shows that are already broadcast). However, this model does not cover scenarios where the content is not fully available at the servers. This situation arises in live streaming applications such as sporting events, where the content is being generated as it is streamed to the end-users. Another relevant application is P2P streaming applications, where the upstream peer has not completely received the content. We would like to emphasize that our notion of live streaming applications are inherently different from video conferencing applications where the streaming is extremely delay-sensitive due to its interactive nature.

In live streaming scenarios, the end-user can generally tolerate some lag from the actual timing of the live event. This lag can be significantly larger than the average

packet delivery delay (round trip time) as well as initial waiting time for buffering. In this model, by causality of the system, the receiver's queue length cannot exceed this lag [2]. Figure 2-6 plots a sample path of the receiver's buffer over time. We are interested in characterizing delay-rate-reliability trade-offs for QoE metrics defined in terms of initial waiting time and interruptions, as a function of the system parameters.

Let us consider the case of Poisson arrivals and a large media file. Note that in this case, even if the arrival rate is larger than the playback rate, the receiver's buffer does not have a positive drift because of the maximum allowable buffer size. Since there is a positive probability of not having any arrivals for a finite period of time, we may easily show that the interruption probability goes to 1 as the file size grows. Therefore, interruption probability is too strict as the QoE metric in live streaming scenarios. However, we may use weaker metrics that capture the transient characteristics of media playback in terms of interruptions, but are not as strict as interruption probability.

2.4.1 Refined Interruption-related QoE Metrics

The probability of interruption as QoE metrics is less restrictive than having per packet deadlines, but quite strong in the sense that having *any* interruptions in the whole playback duration is assumed to be unacceptable. Using this metric, the system is designed so that interruptions are rare events. However, if such events occur e.g. in a live streaming scenario, we need to have a policy to buffer a few blocks of the media file before resuming the playback.

We propose user experience metrics based on end-user related penalties for the gaps in media playback, which facilitates design of re-buffering policies, so that the total end-user penalty regarding the interruptions is minimized. In order to capture the transient behavior of the system, the penalty that we assign to each gap in media playback must be a non-linear function of the duration of the gap. Otherwise, we cannot distinguish many small gaps (which is intuitively less desirable) from a few larger gaps. On the other hand, when the duration of each gap increases beyond

[2]We still assume the playback rate is normalized to one.

some threshold (user's patience), the penalty function should grow at a faster rate. A conceptual example of a candidate for such penalty function is plotted in Figure 2-7.

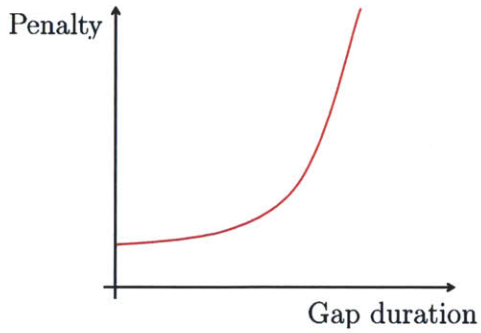

Figure 2-7: The end-user experience degrades as a nonlinear function of the duration of a gap in media playback.

We may consider different scenarios for designing re-buffering policies. The simplest of all is the case with infinite file size, with a memoryless channel where we expect to obtain a stationary control policy. Upon any interruption event, such a policy observes the arrivals and decides to resume the playback based on the number of arrivals and elapsed time from the occurrence of the interruption event. A more challenging case is the scenario of streaming a finite file size. We can formulate this problem as a finite horizon dynamic program. Here, the optimal control policy also depends on the time of the interruption event. Intuitively, as the interruption event occurs closer to the end of the file, the optimal policy requires to re-buffer fewer packets before resuming the playback.

Next, we numerically obtain the optimal trade-off curve between the interruption probability and initial buffer size, and compare the results with the bounds derived earlier.

2.5 Numerical Results

We use MATLAB simulations to compute the minimum initial buffer size $D^*(\epsilon)$ for a given interruption probability ϵ in various scenarios. We start from a small initial

buffer size D, and for each D we compute the interruption probability $p(D)$ via the Monte-Carlo method. We increase D until the constraint $p(D) \leq \epsilon$ is satisfied. Since $p(D)$ is monotonically decreasing in D, this gives the minimum required buffer size. We restrict D to take only integer values, and round each upper bound value up, and each lower bound value down to the nearest integer.

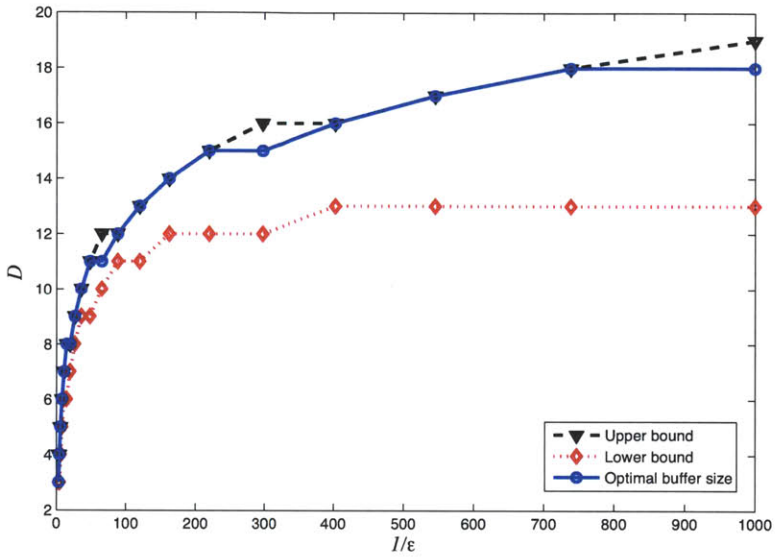

Figure 2-8: The minimum buffer size $D^*(\epsilon)$ as a function of the interruption probability.

Figure 2-8 shows the minimum required buffer size $D^*(\epsilon)$ as a function of $\frac{1}{\epsilon}$, as well as the bounds given by Theorems 2.2.1, 2.2.2. The arrival rate is fixed to $R = 1.2$ and the file size $F = 500$. We observe that the numerically computed trade-off curve closely matches our analytical results.

Figure 2-9 plots the minimum required buffer size $D^*(\epsilon)$ as well as the upper and lower bounds given by Theorems 2.2.1 and 2.2.2 versus the arrival rate R, where $\epsilon = 10^{-2}$ and the file size is fixed to $F = 10^3$. Note that when the arrival rate is almost equal or less than the playback rate, increasing the arrival rate can significantly reduce the initial buffering delay. However, for larger arrival rates $D^*(\epsilon)$ is small enough such that increasing R does not help anymore. Such sensitivity analysis has important practical ramifications. This allows the service provider to understand the actual effect of resource provisioning on the end-user's quality of experience.

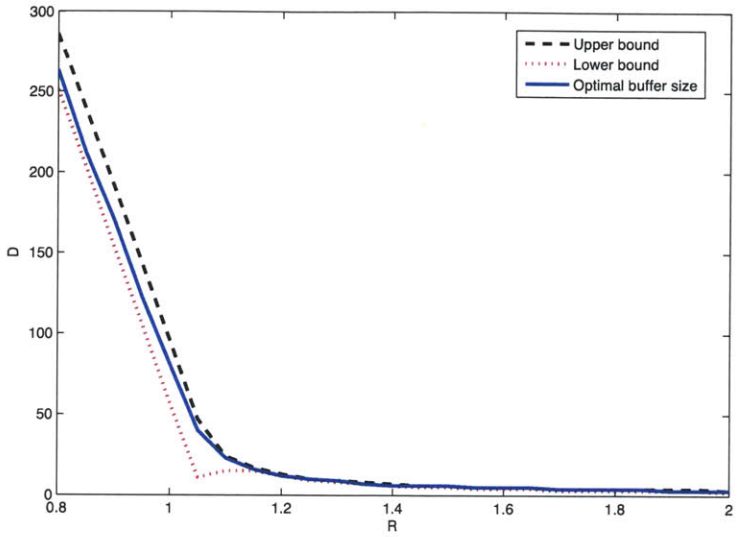

Figure 2-9: The minimum buffer size $D^*(\epsilon)$ as a function of the arrival rate R.

2.6 Appendix to Chapter 2 - Proofs

2.6.1 Proof of the Achievability Theorem

We first prove some useful lemmas used in the proof of Theorem 2.2.1.

Lemma 2.6.1. *Let $I(R)$ be given by (2.5). The following relations hold:*

$$I(R) = 0, \qquad \text{if } 0 \le R \le 1, \tag{2.19}$$

$$\frac{2(R-1)}{R} \le I(R) \le 2(R-1), \qquad \text{if } 1 \le R \le 2, \tag{2.20}$$

$$R - 1 \le I(R) \le R \le 2(R-1), \qquad \text{if } R \ge 2. \tag{2.21}$$

Proof. We consider three different cases separately.

Case I ($0 \le R \le 1$): First note that $\gamma(r)$ is a continuously differentiable function, and $\gamma(0) = 0$. For each $R < 1$, we have $\gamma'(r) > 0$ for all $r \ge 0$. Therefore, $\gamma(r) > 0$ for all $r > 0$, i.e., $I(R) = 0$ for each $R < 1$.

Case II ($1 \le R \le 2$): By definition of $I(R)$ in (2.5),

$$0 = \gamma(I(R)) \;\; = \;\; I(R) + R(e^{-I(R)} - 1) \le I(R) + R(-I(R) + \frac{I^2(R)}{2}).$$

Rearranging the terms in the above relation, gives the lower bound in (2.20). We show the upper bound in two steps. First, we show that $\gamma(2(R-1)) > 0$ for $R > 1$, then we verify that $\gamma(r) \geq 0$ for all $r \geq 2(R-1)$. These two facts imply that $\gamma(r) > 0$ for all $r \geq 2(R-1)$, i.e., $I(R) \leq 2(R-1)$. The first step can be verified by noting that

$$\gamma(2(R-1))\big|_{R=1} = 0, \quad \frac{\partial}{\partial R}\gamma(2(R-1)) > 0.$$

It is also straightforward to show that $\frac{\partial}{\partial r}\gamma(r) > 0$, for all $r \geq \log(R)$, which immediately yields the second step by noting $r \geq 2(R-1) \geq \log(R)$.

Case III ($R \geq 2$): We use a similar technique as in the preceding case. The upper bound is immediate by the following facts:

$$\gamma(R) = Re^{-R} > 0, \quad \frac{\partial}{\partial r}\gamma(r) > 0, \quad \text{for all } r \geq R.$$

We may also check that $\gamma(R-1) < 0$ for all $R \geq 2$. Moreover, note that $\gamma(R) > 0$. Therefore, by the intermediate value theorem, $\gamma(r)$ has a root in $[R-1, R]$, i.e., $I(R) \geq R-1$. □

Lemma 2.6.2. *Let $X(t) = e^{-rQ(t)}$, where $Q(t)$ is given by (3.1). Also let $\gamma(r) = r + R(e^{-r} - 1)$. Then for every $r \geq 0$ such that $\gamma(r) \geq 0$, $X(t)$ is a sub-martingale with respect to the canonical filtration $\mathcal{F}_t = \sigma(X(s), 0 \leq s \leq t)$, i.e., the smallest σ-field containing the history of the stochastic process X up to time t. Moreover, if $\gamma(r) = 0$ then $X(t)$ is a martingale.*

Proof. For every t, $|X(t)| \leq 1$. Hence, $X(t)$ is uniformly integrable. It remains to show that for every $t \geq 0$ and $h > 0$,

$$\mathbf{E}[X(t+h)|\mathcal{F}_t] \geq X(t) \quad \text{a.s.} \tag{2.22}$$

$X(t)$ is a martingale if (2.22) holds with equality. The left-hand side of (2.22) can be

expressed as

$$\begin{aligned}
\mathbf{E}[X(t+h)|\mathcal{F}_t] &= \mathbf{E}\Big[e^{-r(Q(t+h)-Q(t))}\Big|\mathcal{F}_t\Big]X(t) \\
&= \mathbf{E}\Big[e^{-r(A(t+h)-A(t))}\Big|\mathcal{F}_t\Big]e^{rh}X(t) \\
&\overset{(a)}{=} \mathbf{E}\big[e^{-rA(h)}\big]e^{rh}X(t) \\
&\overset{(b)}{=} e^{h(r+R(e^{-r}-1))}X(t) = e^{h\gamma(r)}X(t),
\end{aligned}$$

where (a) follows from independent increment property of the Poisson process, and (b) follows from the fact that $A(t)$ is a Poisson random variable. Now, it is immediate to verify (2.22) for any r with $\gamma(r) \geq 0$. Finally, note that if $\gamma(r) = 0$, the equality in the above relations hold through, and (2.22) holds with equality. Therefore, $X(t)$ is a martingale for r with $\gamma(r) = 0$. □

Next, we use Doob's maximal inequality [44] to bound the interruption probability.

Lemma 2.6.3. *Let $p(D)$ be the interruption probability defined in (2.3), and $\gamma(r) = r + R(e^{-r} - 1)$. Then, for any $r \geq 0$ with $\gamma(r) \geq 0$*

$$p(D) \leq e^{-rD+T\gamma(r)}, \quad \textit{for all } D, T, R \geq 0. \tag{2.23}$$

Proof. By definition of $p(D)$ in (2.3), we have

$$\begin{aligned}
p(D) &= \mathbf{Pr}\{\tau_e < \tau_f\} \leq \mathbf{Pr}\{\tau_e \leq T\} \\
&= \mathbf{Pr}\Big\{\inf_{0\leq t\leq T} Q(t) \leq 0\Big\} \\
&= \mathbf{Pr}\Big\{\sup_{0\leq t\leq T} e^{-rQ(t)} \geq 1\Big\} \\
&\overset{(a)}{\leq} \mathbf{E}[e^{-rQ(T)}] = \mathbf{E}[e^{-r(D+A(T)-T)}] \\
&= e^{-r(D-T)}e^{RT(e^{-r}-1)} = e^{-rD+T\gamma(r)},
\end{aligned}$$

where (a) holds by applying Doob's maximal inequality [44] to the non-negative sub-martingale $X(t) = e^{-rQ(t)}$. Note that $X(t)$ is a sub-martingale for all r with $\gamma(r) \geq 0$ by Lemma 2.6.2. □

Lemma 2.6.4. *It holds that* $-(1-z)\log(1-z) - z \leq -\frac{z^2}{2}$, *for all* $0 \leq z < 1$.

Proof. Let $f(z) = -(1-z)\log(1-z) - z + \frac{z^2}{2}$. $f(z)$ is a continuously differentiable function on $[0,1)$. Moreover, $f(0) = 0$, and $f'(z) = \log(1-z) + z \leq 0$. Therefore, $f(z) \leq f(0) = 0$, for all $z \in [0,1)$. □

Proof of Theorem 2.2.1. First, note that for any upper bound $\bar{p}(D)$ of the interruption probability $p(D)$, any feasible solution of

$$\bar{D}(\epsilon) = \min\{D \geq 0 : \bar{p}(D) \leq \epsilon\} \tag{2.24}$$

provides an upper bound on $D^*(\epsilon)$. This is so since the optimal solution of the above problem is feasible in the minimization problem (2.4). If the problem in (2.24) is infeasible, we use the convention $\bar{D}(\epsilon) = \infty$, which is a trivial bound on $D^*(\epsilon)$. The rest of the proof involves finding the tightest bounds on $p(D)$ and solving (2.24).

Part (a): By Lemma 2.6.3, for $r = I(R)$, we can write $p(D) \leq \bar{p}_a(D) = e^{-I(R)D}$, for all $D, T, R \geq 0$. Solving $\bar{p}_a(D) = \epsilon$ for D gives the result of part (a). Since $I(R) = 0$ for $R \leq 1$ (cf. Lemma 2.6.1), this bound is not useful in that range.

Part (b): First, we claim that for all $D \geq T(1-R+I(R))$, we have $p(D) \leq \bar{p}_b(D) = e^{-\frac{1}{2}TRz^2}$, where $z = 1-\frac{1}{R}\left(1-\frac{D}{T}\right)$. We use Lemma 2.6.3 with $r = r^* = -\log\left(\frac{1}{R}\left(1-\frac{D}{T}\right)\right)$ to prove the claim. Note that $r^* \geq 0$, because $D \geq T(1-R)$. In order to verify the second hypothesis of Lemma 2.6.3, consider the following

$$\begin{aligned} R(e^{-r^*} - e^{-I(R)}) &= I(R) + R(e^{-r^*} - 1) - \gamma(I(R)) \\ &\overset{(a)}{=} I(R) - R + (1 - \frac{D}{T}) \\ &\overset{(b)}{=} \frac{1}{T}\Big[T(1 - R + I(R)) - D\Big] \overset{(c)}{\leq} 0, \end{aligned}$$

where (a) and (b) follow from the definition of $I(R)$ and r^*, respectively, and (c) holds by the hypothesis of the claim. Thus, $r^* \geq I(R)$. Using the facts that $I(R)$ is the largest root of $\gamma(r)$, and $\gamma(r) \to +\infty$ as $r \to \infty$, we conclude that $\gamma(r^*) \geq 0$. Now,

we apply Lemma 2.6.3 to get

$$\begin{aligned} p(D) &\leq \exp\Big(-r^*D + T\gamma(r^*)\Big) \\ &\overset{(a)}{=} \exp\Big(TR\big(\frac{1}{R}(1-\frac{D}{T})r^* - (1-e^{-r^*})\big)\Big) \\ &\overset{(b)}{=} \exp\Big(TR\big(-(1-z)\log(1-z) - z\big)\Big) \\ &\overset{(c)}{\leq} \exp\Big(-\frac{1}{2}TRz^2\Big), \end{aligned}$$

where (a) and (b) follow from the definition of $\gamma(r)$ and z, and (c) is true by Lemma 2.6.4. Therefore, the claim holds.

Now, let $\bar{D} = T(1-R) + \Big(2TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}$. Using the claim that we just proved, we may verify that $p(\bar{D}) \leq \bar{p}_b(\bar{D}) = \epsilon$, if $\bar{D} \geq T(1-R+I(R))$. In order to check the hypothesis of the claim, note that for $R \leq 1$, $I(R) = 0$ (cf. Lemma 2.6.1), and for all $1 \leq R \leq 1 + \Big(\frac{1}{2T}\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}$,

$$\bar{D} - T(1-R) = \Big(2TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}} \geq 2T\Big(\frac{1}{2T}\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}} \overset{(d)}{\geq} 2T(R-1) \overset{(e)}{\geq} TI(R),$$

where (d) follows from the hypothesis, and inequality (e) is true by Lemma 2.6.1. Therefore, $D^*(\epsilon) \leq \bar{D}$ for all $R \leq 1 + \Big(\frac{1}{2T}\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}$. Note that, the upper bound that we obtained in Part (a) is also valid for all R. Hence, the minimum of the two gives the tightest bound. ■

2.6.2 Proof of the Converse Theorem

Next, we establish some lemmas used in the proof of Theorem 2.2.2. The proofs of Lemmas 2.6.5 and 2.6.6 are based on the results from [45], [46].

Lemma 2.6.5. *Let Z be a Poisson random variable with mean λ. If $\lambda \geq 2$, and $k \geq 2$, then*

$$\mathbf{Pr}\{Z \leq \lambda - k\} \leq \exp\Big(\frac{1}{2\lambda}\big(k - \frac{3}{2}\big)^2\Big). \tag{2.25}$$

Proof. Glynn [45] proves that for $\lambda \geq 2$ and $k \geq 2$

$$\mathbf{Pr}\{Z \leq \lambda - k\} \leq b_\lambda Q\Big(\frac{k - \frac{3}{2}}{\sqrt{\lambda}}\Big), \tag{2.26}$$

where

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \quad b_\lambda = (1 + \frac{1}{\lambda})\, e^{\frac{1}{8\lambda}}.$$

The result follows from the facts that $b_\lambda \leq 2$ for $\lambda \geq 2$, and $Q(x) \leq \frac{1}{2} \exp\big(-\frac{x^2}{2}\big)$, for all $x \geq 0$. □

Lemma 2.6.6. *Let Z be a Poisson random variable with mean λ. For all $m \leq \frac{\sqrt{\lambda}}{20} - 1$*

$$\mathbf{Pr}\{Z \leq \lambda - m\sqrt{\lambda}\} \geq \frac{1}{3} \exp\Big(-\frac{1}{1.9}(m + \frac{1}{2})^2\Big). \tag{2.27}$$

Proof. It follows from Proposition 6 of [46] that for all $0 \leq i \leq \lambda/4$

$$\mathbf{Pr}\{Z = \lambda - i\} \geq c_\lambda \exp\Big(-\frac{i^2}{1.9\lambda}\Big), \tag{2.28}$$

where $c_\lambda = \frac{e^{-\frac{1}{12}}}{\sqrt{2\pi\lambda}}$. Therefore, for $m \leq \sqrt{\lambda}/20 - 1$, we may verify that

$$\begin{aligned}
\mathbf{Pr}\{Z \leq \lambda - m\sqrt{\lambda}\} &\geq c_\lambda \sum_{j=0}^{\sqrt{\lambda}-1} \exp\Big(-\frac{(m\sqrt{\lambda} + j)^2}{1.9\lambda}\Big) \\
&\geq c_\lambda e^{-\frac{m^2}{1.9}} \sum_{j=0}^{\sqrt{\lambda}-1} \exp\Big(-\frac{3m}{1.9\sqrt{\lambda}} j\Big) \\
&= c_\lambda e^{-\frac{m^2}{1.9}} \frac{1 - e^{-\frac{3m}{1.9}}}{1 - e^{-\frac{3m}{1.9\sqrt{\lambda}}}} \\
&\geq \Big(\frac{1.9 e^{-\frac{1}{12}}}{3m\sqrt{2\pi}}(1 - e^{-\frac{3m}{1.9}})\Big) e^{-\frac{m^2}{1.9}} \\
&\geq \Big(\frac{1}{3} e^{-\frac{m}{1.9} - \frac{1}{7.6}}\Big) e^{-\frac{m^2}{1.9}} \\
&= \frac{1}{3} \exp\Big(-\frac{1}{1.9}(m + \frac{1}{2})^2\Big). \quad \square
\end{aligned}$$

Lemma 2.6.7. *Let $Q(t)$ be given by (3.1). For $s, \delta > 0$, define the boundary function*

$$u(t) = D + \delta + (R + s - 1)t.$$

If $s \leq R$, then the probability of crossing the boundary is bounded from above as

$$\mathbf{Pr}\Big\{ \sup_{0 \leq t \leq T} Q(t) \geq u(t) \Big\} \leq \exp\big(- \frac{s \cdot \delta}{R} \big).$$

Proof. Define $Z(t) = e^{rQ'(t)}$, where $Q'(t) = Q(t) - u(t) = -\delta + A(t) + (R + s)t$, and $r > 0$ satisfies $\varphi(r) := e^r - 1 - r(1 + \frac{s}{R}) = 0$. Similarly to the proof of Lemma 2.6.2, we can show that $Z(t)$ is a martingale. This allows us to use Doob's maximal inequality to obtain

$$\mathbf{Pr}\Big\{ \sup_{0 \leq t \leq T} Q(t) \geq u(t) \Big\} \;=\; \mathbf{Pr}\Big\{ \sup_{0 \leq t \leq T} Z(t) \geq 1 \Big\} \leq \mathbf{E}[Z(T)] = e^{-r \cdot \delta}.$$

Now it is sufficient to show that $r \geq \frac{s}{R}$. Observe that for all $0 \leq x \leq 1$, $e^x \leq 1 + x + x^2$. Hence, for all $s \leq R$

$$\begin{aligned} \varphi(\frac{s}{R}) &= e^{\frac{s}{R}} - 1 - \frac{s}{R}(1 + \frac{s}{R}) \\ &\leq \frac{s}{R} + (\frac{s}{R})^2 - \frac{s}{R}(1 + \frac{s}{R}) = 0. \end{aligned} \tag{2.29}$$

Moreover, $\varphi(r) \to \infty$ when $r \to \infty$. Therefore, by intermediate value theorem there exists $r \geq \frac{s}{R}$ such that $\varphi(r) = 0$. This completes the proof. □

Proof of Theorem 2.2.2. *Part(a):* Similarly to the argument as in the proof of Theorem 2.2.1, it is sufficient to provide a lower bound on $p(D)$ defined in (2.3).

Define τ_B as the first time that $Q(t)$ crosses a threshold $B > D$, i.e., $\tau_B = \inf\{t : Q(t) \geq B\}$.

A necessary condition for the interruption event to happen is to have the receiver's buffer emptied before time $T' = T - B$, or crossing the threshold B (see Figure 2-10). In particular,

$$p(D) = \mathbf{Pr}\{\tau_e < \tau_f\} \geq \mathbf{Pr}\big\{\tau_e \leq \min\{\tau_B, T'\}\big\}. \tag{2.30}$$

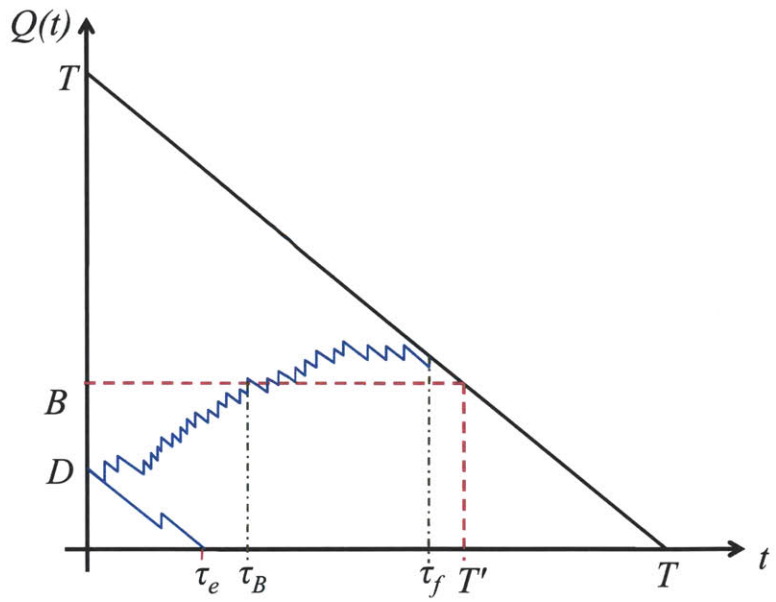

Figure 2-10: Two sample paths of the buffer size $Q(t)$ demonstrating the interruption event at time τ_e, crossing the threshold B at time τ_B, and the download complete event at time τ_f.

Define the stopping time $\tau = \min\{\tau_e, \tau_B, T'\}$, and let $Y(t) = e^{-I(R)Q(t)}$, where $I(R) > 0$ is given by (2.5). By Lemma 2.6.2, $Y(t)$ is a martingale. Moreover, $Y(t) \leq 1$, and $\tau \leq T < \infty$. Therefore, we can apply Doob's optional stopping theorem [44] to get

$$\begin{aligned} e^{-I(R)\cdot D} &= \mathbf{E}[Y(0)] \stackrel{(a)}{=} \mathbf{E}[Y(\tau)] \\ &\stackrel{(b)}{\leq} e^{-I(R)\cdot 0}\big(\mathbf{Pr}\{\tau = \tau_e\} + \mathbf{Pr}\{\tau = T'\}\big) \\ &\quad + e^{-I(R)\cdot B}\big(1 - \mathbf{Pr}\{\tau = \tau_e\} - \mathbf{Pr}\{\tau = T'\}\big) \\ &\leq \mathbf{Pr}\{\tau = \tau_e\} + \mathbf{Pr}\{\tau = T'\} + e^{-I(R)\cdot B} \\ &\stackrel{(c)}{\leq} p(D) + \mathbf{Pr}\{\tau = T'\} + e^{-I(R)\cdot B}, \end{aligned}$$

where (a) is the result of Doob's optional stopping time. (b) holds because $Y(t) \leq 1$ for all t, and $Y(t) \leq e^{-I(R)\cdot B}$ if $Q(t) \geq B$. Finally, (c) follows from (2.30). Rearranging the terms in the above relation, we obtain

$$p(D) \geq e^{-I(R)D} - e^{-I(R)B} - \mathbf{Pr}\{\tau = T'\}. \tag{2.31}$$

Now, choose $B = (1-\alpha)T$, where $\alpha = \frac{R+1}{2R} > \frac{1}{R}$ for all $R > 1$. For all $D, T \geq 2$, we have

$$\begin{aligned}
\mathbf{Pr}\{\tau = T'\} &\overset{(a)}{\leq} \mathbf{Pr}\{0 \leq Q(T-B) \leq B\} \\
&\overset{(b)}{=} \mathbf{Pr}\{\alpha T - D \leq A(\alpha T) \leq T - D\} \\
&\leq \mathbf{Pr}\big\{A(\alpha T) \leq R\alpha T - ((R\alpha - 1)T + D)\big\} \\
&\overset{(c)}{\leq} \exp\bigg(-\frac{1}{2R\alpha T}\big((R\alpha - 1)T + D - \frac{3}{2}\big)^2\bigg) \\
&\overset{(d)}{\leq} \exp\bigg(-\frac{(R-1)^2}{4(R+1)}T\bigg),
\end{aligned} \tag{2.32}$$

where (a) holds because $Q(t)$ cannot be negative or above the threshold B if stopping at T', and (b) follows from the buffer dynamics in (3.1). Recall that $A(\alpha T)$ is a Poisson random variable with mean $R\alpha T$. Since $\alpha \geq \frac{1}{R}$, (c) holds for $D, T \geq 2$ by employing Lemma 2.6.5 with

$$\lambda = R\alpha T \geq T \geq 2, \quad k = (R\alpha - 1)T + D \geq D \geq 2.$$

Finally, (d) is immediate by definition of α noting that $D \geq 2$.

By Lemma 2.6.1, we have $I(R) \geq (R-1)$ for all $R > 1$. Therefore, we can bound the second term in (2.31) as follows

$$e^{-I(R)B} \leq \exp\bigg(-(R-1)\frac{R-1}{2R}T\bigg) \leq \exp\bigg(-\frac{(R-1)^2}{4(R+1)}T\bigg). \tag{2.33}$$

Combine the bounds in (2.32) and (2.33) with (2.31) to obtain

$$p(D) \geq \exp\Big(-I(R)D\Big) - 2\exp\Big(-\frac{(R-1)^2}{4(R+1)}T\Big), \quad \text{for all } D, T \geq 2. \tag{2.34}$$

Therefore, $p(D) \geq \epsilon$ if $D = -\frac{1}{I(R)}\log\Big(\epsilon + 2e^{-\frac{(R-1)^2}{4(R+1)}T}\Big) \geq 2$. This immediately gives the result in (2.8). For the case in which $D < 2$ or $T < 2$, the claim holds trivially.

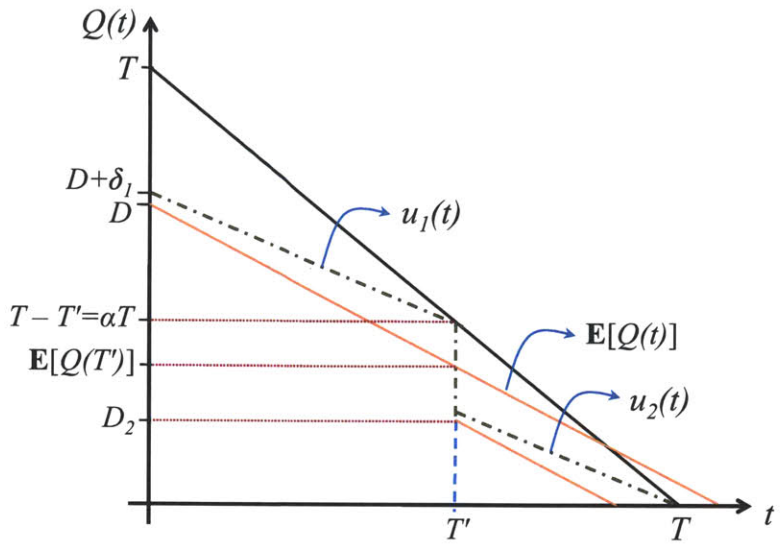

Figure 2-11: Guideline for the proof of Theorem 2.2.2(b).

Part(b): It is sufficient to show $p(D) \geq \epsilon$ for $\epsilon \leq \frac{1}{16}$ and T large enough where

$$D = T(1-R) + \frac{1}{2}\Big(2TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}. \tag{2.35}$$

Let us first define the boundary functions

$$\begin{aligned} u_1(t) &= D + \delta_1 + (R + \frac{\delta_1}{T'} - 1)t, \ t \in [0, T'], \\ u_2(t) &= D_2 + \delta_2 + (R + \frac{\delta_2}{T - T'} - 1)(t - T'), \ t \in [T', T], \end{aligned}$$

where $T' = (1-\alpha)T$ for some constant $0 < \alpha < 1$, and

$$\delta_1 = \frac{1}{2}\alpha TR - \frac{1}{4}\Big(2TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}, \quad \delta_2 = \Big(\alpha TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}, \tag{2.36}$$

$$D_2 = \alpha T(1-R) - 2\Big(\alpha TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}. \tag{2.37}$$

Denote by τ_i the first time to hit the boundary $u_i(t)$, i.e., $\tau_i = \inf\{t : Q(t) \geq u_i(t)\}$, for $i = 1, 2$.

Observe that for every sample path of the buffer size $Q(t)$, the only way for completing the file download, $Q(t) = T - t$, is to cross the boundary functions $u_i(t)$

or have $Q(T') \geq D_2$ (cf. Figure 2-11). This gives a necessary condition for the interruption event. Hence,

$$\begin{aligned}
p(D) &= \mathbf{Pr}\{\tau_e < \tau_f\} \\
&\geq \mathbf{Pr}\{\tau_e \leq \min\{\tau_1, T'\}\} + \mathbf{Pr}\{\tau_e \leq \tau_2, T' \leq \min\{\tau_e, \tau_1\}, Q(T') \leq D_2\} \\
&= \mathbf{Pr}\{\tau_e \leq \min\{\tau_1, T'\}\} + \mathbf{Pr}\{\tau_e \leq \tau_2 | T' \leq \min\{\tau_e, \tau_1\}, Q(T') \leq D_2\}\big[1 \\
&\quad -\mathbf{Pr}\{\tau_1 \leq \min\{\tau_e, T'\}\} - \mathbf{Pr}\{\tau_e \leq \min\{\tau_1, T'\}\} \\
&\quad -\mathbf{Pr}\{T' \leq \min\{\tau_e, \tau_1\}, Q(T') > D_2\}\big] \\
&\geq \mathbf{Pr}\{\tau_e \leq \tau_2 | T' \leq \min\{\tau_e, \tau_1\}, Q(T') \leq D_2\}\big[1 - \mathbf{Pr}\{\tau_1 \leq \min\{\tau_e, T'\}\} \\
&\quad -\mathbf{Pr}\{T' \leq \min\{\tau_e, \tau_1\}, Q(T') > D_2\}\big]. \qquad (2.38)
\end{aligned}$$

In the following we provide bounds on each of the terms in (2.38). By Markov property of the Poisson process we have

$$\begin{aligned}
&\mathbf{Pr}\{\tau_e \leq \tau_2 | T' \leq \min\{\tau_e, \tau_1\}, Q(T') \leq D_2\} \\
&\quad = \mathbf{Pr}\{\tau_e \leq \tau_2 | Q(T') \leq D_2\} \\
&\quad \geq \mathbf{Pr}\{\tau_e \leq \tau_2 | Q(T') = D_2\} \\
&\quad = 1 - \mathbf{Pr}\{\tau_2 \leq \tau_e | Q(T') = D_2\} \\
&\quad \geq 1 - \mathbf{Pr}\Big\{ \sup_{T' \leq t \leq T} Q(t) \geq u_2(t) | Q(T') = D_2 \Big\} \\
&\quad \geq 1 - e^{-\frac{\delta_2^2}{R(T-T')}} = 1 - \epsilon, \qquad (2.39)
\end{aligned}$$

where the last inequality follows from Lemma 2.6.7 with parameters $\delta = \delta_2$ and $s = \frac{\delta_2}{T-T'}$, if its hypothesis $s \leq R$ is satisfied. This is equivalent to having T satisfy $T \geq \frac{1}{\alpha R} \log\left(\frac{1}{\epsilon}\right)$.

Similarly, by employing Lemma 2.6.7 with $\delta = \delta_1$ and $s = \frac{\delta_1}{T'}$ we have

$$\begin{aligned}
\mathbf{Pr}\{\tau_1 \leq \min\{\tau_e, T'\}\} &\leq \mathbf{Pr}\Big\{ \sup_{0\leq t\leq T'} Q(t) \geq u_1(t)\Big\} \\
&\leq \exp\Big(- \frac{\delta_1^2}{RT'}\Big) \\
&= \exp\Big(- \frac{\alpha^2 RT}{4(1-\alpha)}\Big[1 - \Big(\frac{\log\left(\frac{1}{\epsilon}\right)}{2\alpha^2 RT}\Big)^{\frac{1}{2}}\Big]^2\Big).
\end{aligned}$$

Note that the hypothesis of Lemma 2.6.7 is satisfied here for all $\alpha \leq \frac{2}{3}$. Moreover, if $T \geq \frac{16}{\alpha^2 R}\log\left(\frac{1}{\epsilon}\right)$, we have

$$\begin{aligned}
\mathbf{Pr}\{\tau_1 \leq \min\{\tau_e, T'\}\} &\leq \exp\Big(- 4\log\big(\frac{1}{\epsilon}\big)[1 - \frac{1}{\sqrt{32}}]^2\Big) \\
&\leq \exp\Big(- 2\log\big(\frac{1}{\epsilon}\big)\Big) = \epsilon^2. \qquad (2.40)
\end{aligned}$$

For the last term in (2.38) write

$$\begin{aligned}
1 - \mathbf{Pr}\{T' \leq \min\{\tau_e, \tau_1\}, Q(T') > D_2\} &\geq 1 - \mathbf{Pr}\{Q(T') > D_2\} = \mathbf{Pr}\{Q(T') \leq D_2\} \\
&= \mathbf{Pr}\{D + A(T') - T' \leq D_2\} \\
&= \mathbf{Pr}\{A(T') \leq T'R - m\sqrt{T'R}\}, \qquad (2.41)
\end{aligned}$$

where $A(T')$ is a Poisson random variable with mean RT', and

$$m = \Big(\frac{\log\left(\frac{1}{\epsilon}\right)}{2(1-\alpha)}\Big)^{\frac{1}{2}}\big[1 + (8\alpha)^{\frac{1}{2}}\big]. \qquad (2.42)$$

For $T \geq \frac{16}{\alpha^2 R}\log\left(\frac{1}{\epsilon}\right)$ and $\alpha, \epsilon \leq \frac{1}{16}$, we may verify that $m \leq \sqrt{RT'}/20 - 1$. Hence, we can use Lemma 2.6.6 to bound (2.41) from below and conclude

$$1 - \mathbf{Pr}\{T' \leq \min\{\tau_e, \tau_1\}, Q(T') > D_2\} \geq \frac{1}{3}\exp\Big(-\frac{1}{1.9}\big(m + \frac{1}{2}\big)^2\Big).$$

Observe that for $\alpha = 0$, we have $m = m_0 = \left(\frac{1}{2}\log\left(\frac{1}{\epsilon}\right)\right)^{\frac{1}{2}}$ and verify that

$$\frac{1}{1.9}\left(m_0 + \frac{1}{2}\right)^2 < \log\left(\frac{1}{\epsilon}\right) - \log\left(\frac{17}{15}\right), \quad \text{for all } \epsilon \leq \frac{1}{16}.$$

By continuity of m in α (cf. (2.42)), we can choose $\alpha = \alpha_0 > 0$ small enough such that $\frac{1}{3}e^{-\frac{1}{1.9}\left(m+\frac{1}{2}\right)^2} \geq \frac{17}{15}\epsilon$. Now, by plugging this relation as well as (2.39) and (2.40) back in (2.38) we have for all $\epsilon \leq \frac{1}{16}$,

$$p(D) \geq (1-\epsilon)\left(\frac{17}{15}\epsilon - \epsilon^2\right) \geq \frac{15}{16}\left(\frac{17}{15} - \frac{1}{16}\right)\epsilon \geq \epsilon, \tag{2.43}$$

if $T \geq \frac{16}{\alpha_0^2 R}\log\left(\frac{1}{\epsilon}\right)$. Therefore, the buffer size D, defined in (2.35), is a lower bound on $D^*(\epsilon)$. ■

Chapter 3

Network Coding for Technology-heterogeneous Streaming Systems

In this part, we study the problem of streaming a media file from multiple servers to a single receiver over unreliable communication channels. Each of the servers could be a wireless access point, base station, another peer, or any combination of the above. Such servers may operate under different protocols in different ranges of the spectrum such as WiFi (IEEE 802.11), WiMAX (IEEE 802.16), HSPA, EvDo, LTE, etc. We refer to such system as a technology-heterogeneous multi-server system. In this setup, the receiver can request different pieces of the media file from different servers. Requesting packets form each server may cause delays due to channel uncertainty. However, requesting one packet from multiple servers introduces the need to keep track of packets, and the duplicate packet reception problem. In this chapter, we discuss methods that enable efficient streaming across different paths and network interfaces. This greatly simplifies the model when analyzing such systems.

3.1 Multi-path Single-server Systems

In order to resolve such issues of the multi-server and technology-heterogeneous systems, let us take a closer look at the process of media streaming across different layers. Media files are divided into blocks of relatively large size, each consisting of several frames. The video coding is such that all the frames in the block need to be available before any frames can be played. Blocks are requested in sequence by the playback application from the user-end. The server (or other peers) packetize the requested block and transmit them to the user as in Figure 3-1.

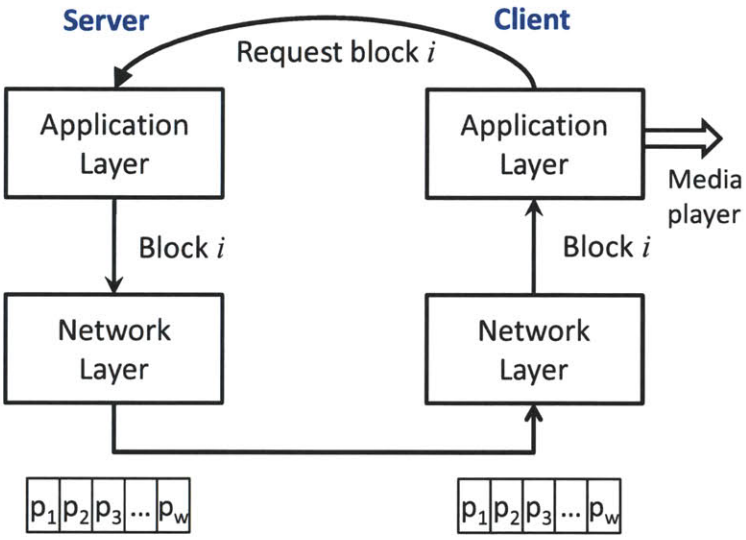

Figure 3-1: The media player (application layer) requires complete blocks. At the network layer each block is divided into packets and delivered.

Now consider the scenario, illustrated in Figure 3-2, where there are multiple paths to reach a particular server. Each of these paths could pass through different network infrastructures. For example, in Figure 3-2, one of the path is using the WiFi access point, while the other one is formed by the Lte network.

The conventional approach in exploiting the path diversity in such scenarios is scheduling each packet to be delivered over one of the available paths. For instance, odd-numbered packets are assigned to path 1, and even-numbered packets are assigned to path 2. This approach requires a deterministic and reliable setting, where each path is lossless and the capacity and end-to-end delay of each path is known.

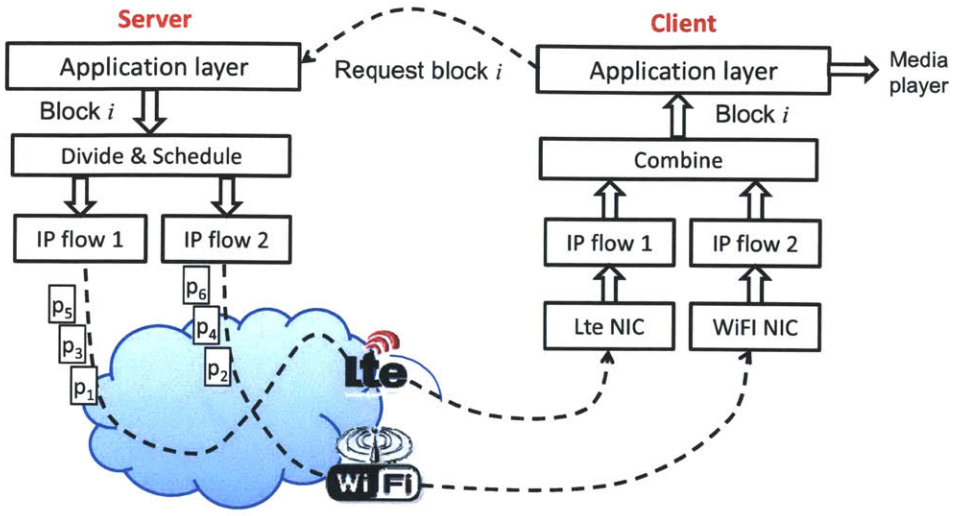

Figure 3-2: Streaming over multiple paths/interfaces.

However, the wireless medium is intrinsically unreliable and time varying. Moreover, flow dynamics in other parts of the network may result in congestion on a particular path. Therefore, the slowest or most unreliable path becomes the bottleneck. In order to compensate for that, the scheduler may add some redundancy by sending the same packet over multiple paths which results in duplicate packet reception and loss of performance. There is a significant effort to use proper scheduling and control mechanisms to reduce these problems. For more information on this approach, generally known as MultiPath TCP (MPTCP), please refer to the works by Wischik *et al.* [47, 48], and IETF working draft on MPTCP [49].

We propose random linear network coding (RLNC) to alleviate the duplicate packet reception problem. Figure 3-3 illustrates an example. Here, instead of requesting a particular packet in block i, the receiver simply requests a random linear combination of all the packets in block i. The coefficients of each combination are chosen uniformly at random from a Galois field of size q. The coded packets delivered to the receiver can be thought of as linear equations, where the unknowns are the original packets in block i. Block i can be fully recovered by solving a system of linear equations if it is full rank. Note that we can embed the coding coefficients in the header of each coded packet so that the receiver can form the system of linear equations. For more implementation details, please refer to Sundararajan *et al.* [50].

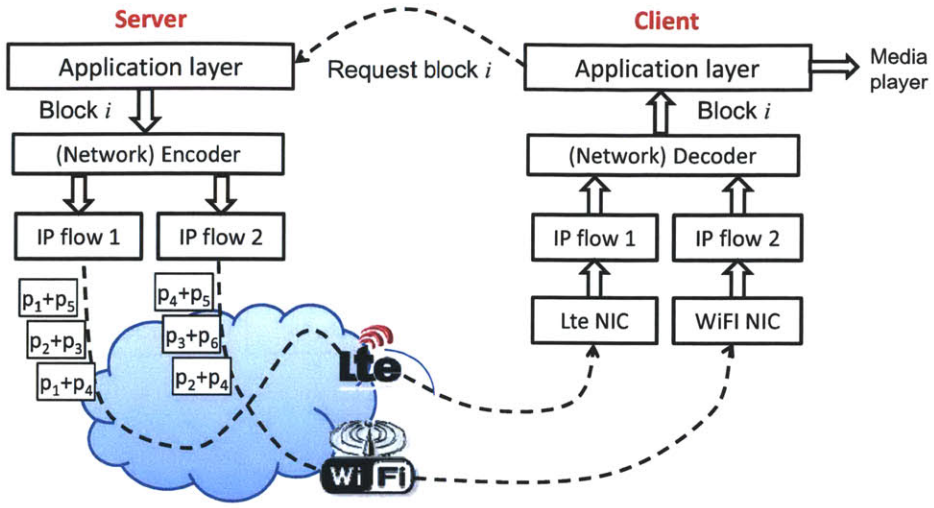

Figure 3-3: Streaming over multiple paths/interfaces using Network Coding.

It can be shown that if the field size q is large enough, the received linear equations are linearly independent with very high probability [24]. Therefore, for recovering a block of W packets, it is sufficient to receive W linearly independent coded packets from different peers. Each received coded packet is likely to be independent of previous ones with probability $1 - \delta(q)$, where $\delta(q) \to 0$ as $q \to \infty$.

By removing the notion of *unique identity* assigned to each packet, network coding allows significant simplification of the scheduling and flow control tasks at the server. For instance, if one of the paths get congested or drops a few of the packets, the server may complete the block transfer by sending more coded packets over the other paths. Hence, the sender may perform TCP-like flow management and congestion control on each of the paths *independently*. Therefore, network coding provides a mean to homogenize a technology-heterogeneous system as if the receiver only has single interface. The following proposition summarizes the corollaries of random linear network coding with a large field size.

Proposition 3.1.1. *Consider a single client receiving random linearly coded packets over m paths with rates $R_1, \ldots R_m$. The effective rate of receiving decoded packets is*

$$R = \sum_{k=1}^{m} R_k.$$

3.2 Multi-path Multi-server Systems

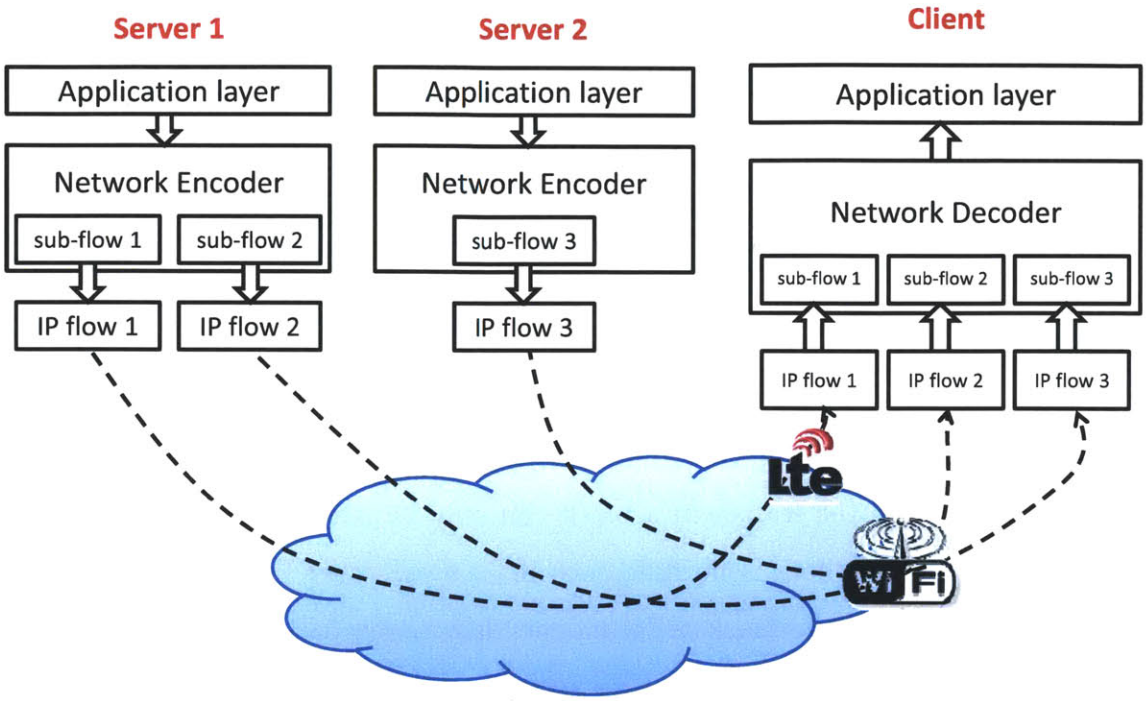

Figure 3-4: Streaming over multiple paths/interfaces from multiple servers/peers using Network Coding.

Our observations in a single-server multi-path case, lead us to further generalization of the multi-path streaming scenario to cases with multiple servers or peers. Figure 3-4 shows a scenario where two independent servers stream a single file over three paths to a client with two interfaces. In this scenario, there are three network (IP) flows corresponding to three paths that are aggregated at the client. Each path delivers random linear combinations of packets in a particular block of the file. The number of coded packets delivered per unit time depends on the level of congestion and the over quality of each path. However, it does not depend on the quality or congestion along other paths.

Figure 3-4 also represents P2P systems, where each of the randomly contacted peers act as a server. In a standard P2P system, it is unlikely that a randomly contacted peer would have all packets corresponding to a particular block. However, storing blocks in a random linear coded fashion ensures that the selected peer has a

useful equation to offer (see [27] for further discussion). Therefore, we may assume that the effective rate at which each path delivers packets to the client is only a function of path state (e.g. congestion and losses) not server state. The following proposition states this observation.

Proposition 3.2.1. *Consider a single client receiving random linearly coded packets over m paths (with independent conditions) from n different servers. The packet arrival process from each path $k \in \{1, \ldots, m\}$ is independent of the arrival process on all other paths.*

Note that such random linear coding does not introduce additional decoding delay for each block, since the frames in a block can only be played out when the whole block is received. So there is no difference in delay whether the end-user received W uncoded packets of the block or W independent coded packets that can then be decoded.

In the following, we discuss the conditions under which we can convert a access technology-heterogeneous multi-server system to a single-path single-server system. Consider a single user receiving a media file from various servers it is connected to. Assume that the media file is divided into blocks of W packets. Each server sends random linear combinations of the packets within the current block to the receiver. We assume that the linear combination coefficients are selected from a Galois field of large enough size, so that no redundant packet is delivered to the receiver. Moreover, we assume that the block size W is small compared to the total length of the file, but large enough to ignore the boundary effects of moving from one block to the next. Assume time is continuous, and the arrival process of packets from each server is a Poisson process. Using network coding, by Proposition 3.2.1, we may assume that the arrival process from each server is independent of other arrival processes. Moreover, since no redundant packet is delivered from different peers (Proposition 3.1.1), we can combine the arrival processes into one Poisson process of the sum-rate R. Similarly to the preceding Chapter, we normalize the playback rate to one, i.e., it takes one unit of time to play a single packet. Thus, our simplified model is just a

single-server-single-receiver system with the queue dynamics given by (3.1).

The presence of some packets in the buffer does not guarantee no interruption, since we require W packets corresponding to a block before it can be decoded. However, if there are at least W packets in the buffer, there is at least one playable packet. This is so since either the first W packets in the buffer belong to the same block, or they belong to two different blocks. In the former case, the packets of the block can be decoded, and in the latter case, the first block of the two must be already decoded; otherwise, the next block would not be sent from the server. We declare an interruption in playback when the buffer size decreases to the threshold W. For simplicity of notation, we assume that an *extra* block is initially buffered (not counted in the initial buffer size D). Hence, we can declare an interruption in playback when the buffer size reaches zero before reaching the end of the file. Since the queue dynamics and the interruption event match the single-server case, the results of the preceding chapter (in particular Theorems 2.2.1, 2.2.2) directly apply to a technology-heterogeneous multi-server system. We summarize the above discussions into the following Proposition, which is the key for development of the analytical results in the subsequent parts.

Proposition 3.2.2. *Consider one or more servers streaming a single media file to a single client over m independent path using random linear network coding. The packet delivery process over path k is modeled as a Poisson process of rate R_k. Let the playback rate be normalized to one at the receiver. Then, the receiver queue dynamics is given by:*

$$Q(t) = D + A(t) - t, \tag{3.1}$$

where D is the initial queue length and $A(t)$ is a Poisson process of rate $R = \sum_{k=1}^{m} R_k$.

We declare an interruption in media playback if and only if the queue length hits zero before reaching the end of the file, i.e., when $\tau_e < \tau_f$, where τ_e and τ_f are defined in (2.2).

Proposition 3.2.2 provides the necessary tool for analyzing technology-heterogeneous multi-server systems as a single-server system. For instance, we can apply most of the

results of Chapter 2 on fundamental delay-interruption trade-offs. This is essential for tractability of the analysis of cost-heterogeneous systems, which is the focus of the subsequent Chapter.

Chapter 4

Cost-heterogeneous Multi-server Streaming Systems

In this part, we study the multi-server media streaming problem, where the cost of accessing different servers can be different. This scenario adds another dimension to end-user metrics, i.e., the usage cost (see Figure 1-3). The goal is to satisfy user's quality of experience requirements in terms of the initial waiting time and interruption probability at a minimum cost.

4.1 System Model and QoE Metrics

We consider a media streaming system as follows. A single user is receiving a media file of size F packets, from various servers or access points. The receiver first buffers D packets from the beginning of the file, and then starts the playback at unit rate.

We assume that time is continuous, and the arrival process of packets from each server is a Poisson process independent of other arrival processes. Further, we assume that each server sends random linear combination of the packets in the source file. Therefore, by discussions of Chapter 3 (cf. Proposition 3.1.1), no redundant packet is delivered from different servers. Therefore, we can combine the arrival processes of any subset of the servers into one Poisson process of rate equal to the sum of the rates from the corresponding servers (cf. Proposition 3.2.2).

There are two types of servers in the system: free servers and the costly ones. There is no cost associated with receiving packets from a free server, but a unit cost is incurred for every unit time that the costly servers are used. As described above, we can combine all the free servers into one free server from which packets arrive according to a Poisson process of rate R_0. Similarly, we can merge all of the costly servers into one costly server with effective rate of R_c. At any time t, the user has the option to use packets only from the free server or from both the free and the costly servers. In the latter case, the packets arrive according to a Poisson process of rate

$$R_1 = R_0 + R_c.$$

The user's action at time t is denoted by $u_t \in \{0, 1\}$, where $u_t = 0$ if only the free server is used at time t, while $u_t = 1$ if both free and costly servers are used. We assume that the parameters R_0 and R_1 are known at the receiver. Figure 4-1 illustrates the system model.

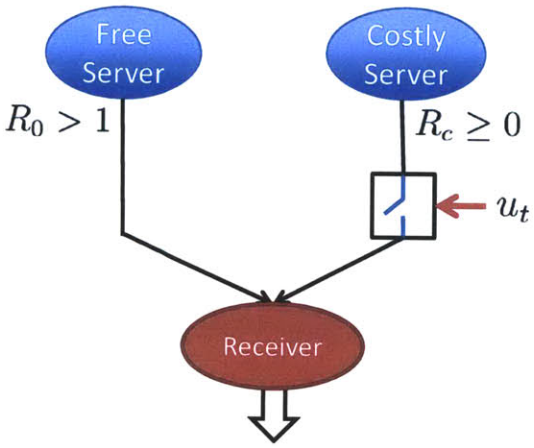

Figure 4-1: Streaming frow two classes of servers: costly and free.

The dynamics of the receiver's buffer size (queue-length) Q_t can be described as follows

$$Q_t = D + N_t + \int_0^t u_\tau dN_\tau^c - t, \tag{4.1}$$

where D is the initial buffer size, N_t Poisson processes of rate R_0 and N_t^c is a Poisson counter of rate R_c which is independent of the process N_t. The last term correspond

to the unit rate of media playback.

The user's association (control[1]) policy is formally defined below.

Definition 4.1.1. [Control Policy] Let

$$h_t = \{Q_s : 0 \leq s \ \leq t\} \cup \{u_s : 0 \leq s \ < t\}$$

denote the history of the buffer sizes and actions up to time t, and $\mathcal{H}$ be the set of all histories for all t. A *deterministic association policy* denoted by π is a mapping $\pi : \mathcal{H} \longmapsto \{0, 1\}$, where at any time t

$$\pi(h_t) = \begin{cases} 0, & \text{if only the free server is chosen,} \\ 1, & \text{if both servers are chosen.} \end{cases}$$

Denote by Π the set of all such control policies.

Similarly to Section 2.1, we use the initial buffer size D, and interruption probability as quality of user experience metrics. The definition of interruption event and hence, interruption probability remains unchanged (see (2.3)). However, the interruption event not only depends on the initial buffer size D, but also on the control policy π. To emphasize this dependency, we denote the interruption probability by

$$p^{\pi}(D) = \mathbf{Pr}\{\tau_e < \tau_f\}, \tag{4.2}$$

where τ_e and τ_f are defined in (2.2).

Definition 4.1.2. The policy π is defined to be (D, ϵ)-*feasible* if $p^{\pi}(D) \leq \epsilon$. The set of all such feasible policies is denoted by $\Pi(D, \epsilon)$.

The third metric that we consider in this thesis is the expected cost of using the costly server which is proportional to the expected usage time of the costly server.

[1]Throughout the rest of this chapter, we use the notion of control policy and association policy, interchangeably.

For any (D, ϵ), the usage cost of a (D, ϵ)-feasible policy π is given by[2]

$$J^{\pi}(D, \epsilon) = \mathbf{E}\Big[\int_0^{\min(\tau_e, \tau_f)} u_t dt\Big]. \tag{4.3}$$

The *value function* or optimal cost function V is defined as

$$V(D, \epsilon) = \min_{\pi \in \Pi(D,\epsilon)} J^{\pi}(D, \epsilon), \tag{4.4}$$

and the optimal policy π^* is defined as the optimal solution of the minimization problem in (4.4).

In our model, the user expects to buffer no more than D packets and have an interruption-free experience with probability higher than a desired level $1 - \epsilon$. Note that there are trade-offs between the interruption probability ϵ, the initial buffer size D, and the usage cost. These trade-offs depend on the association policy as well as the system parameters R_0, R_c and F.

Throughout the rest of this chapter, we study the case that $R_0 > 1$ and the file size F goes to infinity, so that the control policies in this case take simpler forms. Moreover, the cost of such control policies provide an upper bound for the finite file size case.

We first characterize the region of interest in the space of QoE metrics where a feasible control policy exists and is non-degenerate. We then use these results to design association policies.

Theorem 4.1.1. *Let (D, ϵ) be user's QoE requirement when streaming an infinite file from two servers. The arrival rate of the free server is given by $R_0 > 1$, and the total arrival rate when using the costly server is denoted by $R_1 > R_0$. Let $I(R)$ be the largest root of $\gamma(r) = r + R(e^{-r} - 1)$. Then*

(a) For any (D, ϵ) such that $D \geq \frac{1}{I(R_0)} \log\left(\frac{1}{\epsilon}\right)$,

$$\min_{\pi \in \Pi} J^{\pi}(D, \epsilon) = 0.$$

[2]Throughout this work, we use the convention that the cost of an infeasible policy is infinite.

(b) For any (D, ϵ) *such that* $D < \frac{1}{I(R_1)} \log\left(\frac{1}{\epsilon}\right)$,

$$\min_{\pi \in \Pi} J^{\pi}(D, \epsilon) = \infty.$$

Proof. Consider the degenerate policy $\pi_0 \equiv 0$. This policy is equivalent to a single-server system with arrival rate $R = R_0$. By Definition 4.1.2, and Corollary 2.2.2, the policy π_0 is (D, ϵ)-feasible for all $D \geq \frac{1}{I(R_0)} \log\left(\frac{1}{\epsilon}\right)$. Note that by (4.3) this policy does not incur any cost, which results in part (a).

Moreover, for all (D, ϵ) with $D < \frac{1}{I(R_1)} \log\left(\frac{1}{\epsilon}\right)$, there is no (D, ϵ)-feasible policy. This is so since the buffer size under any policy π is stochastically dominated by the one governed by the degenerate policy $\pi_1 \equiv 1$. Hence,

$$p^{\pi}(D) \geq p^{\pi_1}(D) = \exp(-I(R_1)D) > \epsilon.$$

Using the convention of infinite cost for infeasible policies, we obtain the result in part (b). □

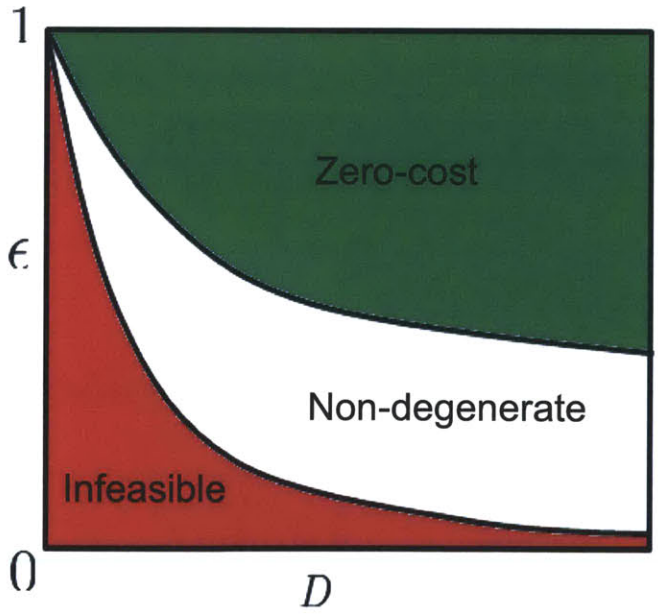

Figure 4-2: Non-degenerate, zero-cost and infeasible regions for QoE metrics (D, ϵ).

For simplicity of notation, let $\alpha_0 = I(R_0)$, and $\alpha_1 = I(R_1)$. By Theorem 4.1.1 we

focus on the region

$$\mathcal{R} = \left\{(D, \epsilon) : \frac{1}{\alpha_1} \log\left(\frac{1}{\epsilon}\right) \leq D \leq \frac{1}{\alpha_0} \log\left(\frac{1}{\epsilon}\right)\right\} \tag{4.5}$$

to analyze the expected cost of various classes of control policies. Figure 4-2 illustrates a conceptual example of this non-degenerate region as well as the zero-cost and infeasible regions.

4.2 Design and Analysis of Association Policies

In this part, we propose several classes of parameterized control policies. We first characterize the range of the parameters for which the association policy is feasible for a given initial buffer size D and the desired level of interruption probability ϵ. Then, we try to choose the parameters such that the expected cost of the policy is minimized. All of the proofs of the main theorems are included in Appendix 4.5.

4.2.1 Off-line Policy

Consider the class of policies where the decisions are made off-line before starting the media streaming. In this case, the arrival process is not observable by the decision maker. Therefore, the user's decision space reduces to the set of deterministic functions $u : \mathbb{R} \to \{0, 1\}$, that maps time into the action space.

Theorem 4.2.1. *Let the cost of a control policy be defined as in (4.3). In order to find a minimum-cost off-line policy, it is sufficient to consider policies of the form:*

$$\pi(h_t) = u_t = \begin{cases} 1, & \text{if } t \leq t_s \\ 0, & \text{if } t > t_s, \end{cases} \tag{4.6}$$

which parameterized are by a single parameter $t_s \geq 0$.

Proof. In general any off-line policy π consists of multiple intervals in which the costly server is used. Consider an alternative policy π' of the form of (4.6) where $t_s = J^\pi$. By

definition of the cost function in (4.3) the two policies incur the same cost. Moreover, the buffer size process under policy π is stochastically dominated by the one under policy π', because the policy π' counts the arrivals from the costly server earlier, and the arrival process is stationary. Hence, the interruption probability of π' is not larger than that of π. Therefore, for any off-line policy, there exists another off-line policy of the form given by (4.6). □

Theorem 4.2.2. *Consider the class of off-lines policies of the form (4.6). For any* $(D, \epsilon) \in \mathcal{R}$, *the policy* π *defined in (4.6) is feasible if*

$$t_s \geq t_s^* = \frac{R_0}{R_1 - R_0}\left[\frac{1}{\alpha_0}\log\left(\frac{1}{\epsilon - e^{-\alpha_1 D}}\right) - D\right]. \tag{4.7}$$

Note that obtaining the optimal off-line policy is equivalent to finding the smallest t_s for which the policy is still feasible. Therefore, t_s^* given in (4.7) provides an upper bound on the minimum cost of an off-line policy. Observe that t_s^* is almost linear in D for all (D, ϵ) that is not too close to the lower boundary of region $\mathcal{R}$. As (D, ϵ) gets closer to the boundary, t_s^* and the expected cost grows to infinity, which is in agreement with Theorem 4.1.1. In this work we pick t_s^* as a benchmark for comparison to other policies that we present next.

4.2.2 Online Safe Policy

Let us now consider the class of online policies where the decision maker can observe the buffer size history. Inspired by the structure of the optimal off-line policies, we first focus on a *safe* control policy in which in order to avoid interruptions, the costly server is used at the beginning until the buffer size reaches a certain threshold after which the costly server is never used. This policy is formally defined below.

Definition 4.2.1. The online safe policy π^S parameterized by the threshold value S is given by

$$\pi^S(h_t) = \begin{cases} 1, & \text{if } t \leq \tau_s \\ 0, & \text{if } t > \tau_s, \end{cases} \tag{4.8}$$

where $\tau_S = \inf\{t \geq 0 : Q_t \geq S\}$.

Theorem 4.2.3. *Let π^S be the safe policy defined in Definition 4.2.1. For any $(D, \epsilon) \in \mathcal{R}$, the safe policy is feasible if*

$$S \geq S^* = \frac{1}{\alpha_0} \log\left(\frac{1}{\epsilon - e^{-\alpha_1 D}}\right). \tag{4.9}$$

Moreover,

$$\min_{S \geq S^*} J^{\pi^S}(D, \epsilon) = J^{\pi^{S^*}}(D, \epsilon) = \frac{1}{R_1 - 1}\left[\frac{1}{\alpha_0} \log\left(\frac{1}{\epsilon - e^{-\alpha_1 D}}\right) - D + \xi\right],$$

where $\xi \in [0, 1)$.

Let us now compare the online safe policy π^{S^*} with the off-line policy defined in (4.6) with parameter t_s^* as in (4.7). We observe that the cost of the online safe policy is almost proportional to that of the off-line policy, where the cost ratio of the off-line policy to that of the online safe policy is given by

$$\frac{R_0(R_1 - 1)}{R_1 - R_0} = 1 + \frac{R_1(R_0 - 1)}{R_1 - R_0} > 1.$$

Note that the structure of both policies is the same, i.e, both policies use the costly server for a certain period of time and then switch back to the free server. As suggested here, the advantage of observing the buffer size allows the online policies to avoid excessive use of the costly server when there are sufficiently large number of arrivals from the free server. In the following, we present another class of online policies.

4.2.3 Online Risky Policy

In this part, we study a class of online policies where the costly server is used only if the buffer size is below a certain threshold. We call such policies "risky" as the risk of interruption is spread out across the whole trajectory unlike the "safe" policies. Further, we constrain risky policies to possess the property that the action at a particular time should only depend on the buffer size at that time, i.e., such policies

are *stationary Markov* with respect to buffer size as the state of the system. The risky policy is formally defined below.

Definition 4.2.2. The online risky policy π^T parameterized by the threshold value T is given by

$$\pi^T(h_t) = \pi^T(Q_t) = \begin{cases} 1, & \text{if } 0 < Q_t < T \\ 0, & \text{otherwise.} \end{cases} \tag{4.10}$$

Theorem 4.2.4. *Let π^T be the risky policy defined in Definition 4.2.2. For any $(D, \epsilon) \in \mathcal{R}$, the policy π^T is feasible if the threshold T satisfies*

$$T \geq T^* = \begin{cases} \frac{1}{\alpha_1 - \alpha_0}\left[\log\left(\frac{\beta}{\epsilon}\right) - \alpha_0 D\right], & \text{if } D \geq \bar{D}, \\ \frac{1}{\alpha_1}\log\left(\frac{\epsilon + \beta(1 - e^{-\alpha_1 D}) - 1}{\epsilon - e^{-\alpha_1 D}}\right), & \text{if } D \leq \bar{D}, \end{cases} \tag{4.11}$$

where $\beta = \frac{\alpha_1}{\alpha_0(1 - \frac{\alpha_0}{2})}$ and $\bar{D} = \frac{1}{\alpha_1}\log\left(\frac{\beta}{\epsilon}\right)$.

Theorem 4.2.4 facilitates the design of risky policies with a single-threshold structure, for any desired initial buffer size D and interruption probability ϵ. For a fixed ϵ, when D increases, T^* (the design given by Theorem 4.2.4) decreases to zero. On the other hand, if D decreases to $\frac{1}{\alpha_1}\log\left(\frac{1}{\epsilon}\right)$ (the boundary of $\mathcal{R}$), the threshold T^* quickly increases to infinity, i.e., the policy does not switch back to the free server unless a sufficiently large number of packets is buffered. Figure 4-3 plots T^* and D as a function of D for a fixed ϵ. Observe that for large range of D, $T^* \leq D$, i.e., the costly server is not initially used. In this range, owing to the positive drift of Q_t, the probability of ever using the costly server exponentially decreases in $(D - T^*)$.

Next we compute relatively tight bounds on the expected cost of the online risky policy and compare with the previously proposed policies.

Theorem 4.2.5. *For any $(D, \epsilon) \in \mathcal{R}$, consider an online risky policy π^{T^*} defined in Definition 4.2.2, where the threshold T^* is given by (4.11) as function of D and ϵ. If $D \geq \bar{D}$ then*

$$J^{\pi^{T^*}}(D, \epsilon) \leq \frac{\beta}{\alpha_1(R_1 - 1)} e^{-a_0(D - T^*)}, \tag{4.12}$$

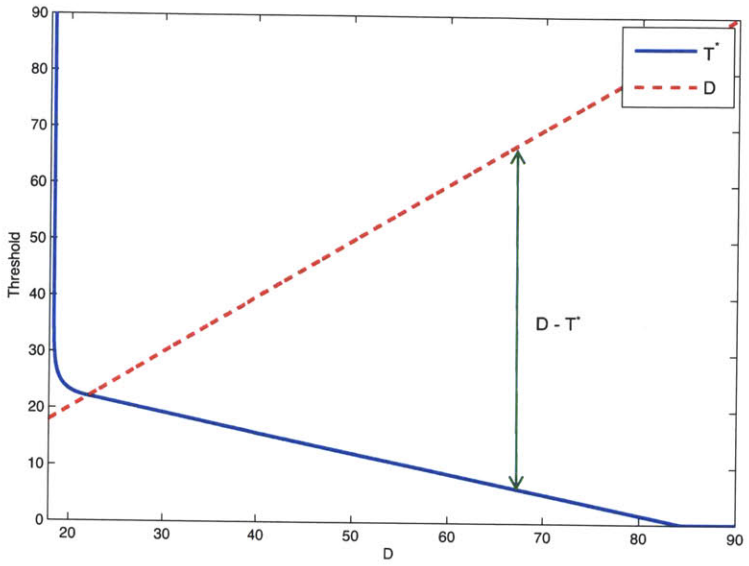

Figure 4-3: The switching threshold of the online risky policy as a function of the initial buffer size for $\epsilon = 10^{-3}$ (See Theorem 4.2.4).

and if $D \leq \bar{D}$

$$J^{\pi^{T^*}}(D, \epsilon) \leq \frac{1 - e^{-\alpha_1 D}}{(R_1 - 1)(1 - e^{-\alpha_1 T^*})}\left(T^* + 1 + \frac{\beta}{\alpha_1}\right) - \frac{D}{R_1 - 1}, \tag{4.13}$$

where $\beta = \frac{\alpha_1}{\alpha_0(1-\frac{\alpha_0}{2})}$ *and* $\bar{D} = \frac{1}{\alpha_1}\log\left(\frac{\beta}{\epsilon}\right)$.

In the following, we compare the expected cost of the presented policies using numerical methods, and illustrate that the bounds derived in Theorems 4.2.2, 4.2.3 and 4.2.5 on the expected cost function are close to the exact value.

4.2.4 Performance Comparison

Figure 4-4 compares the expected cost functions of the off-line, online safe and online risky policies as a function of the initial buffer size D, when the interruption probability is fixed to $\epsilon = 10^{-3}$, the arrival rate from the free server is $R_0 = 1.05$, and the arrival rate from the costly server is $R_c = R_1 - R_0 = 0.15$. We plot the bounds on the expected cost given by Theorems 4.2.2, 4.2.3 and 4.2.5 as well as the expected cost function numerically computed by the Monte-Carlo method.

Observe that the expected cost of the risky policy is significantly smaller that both

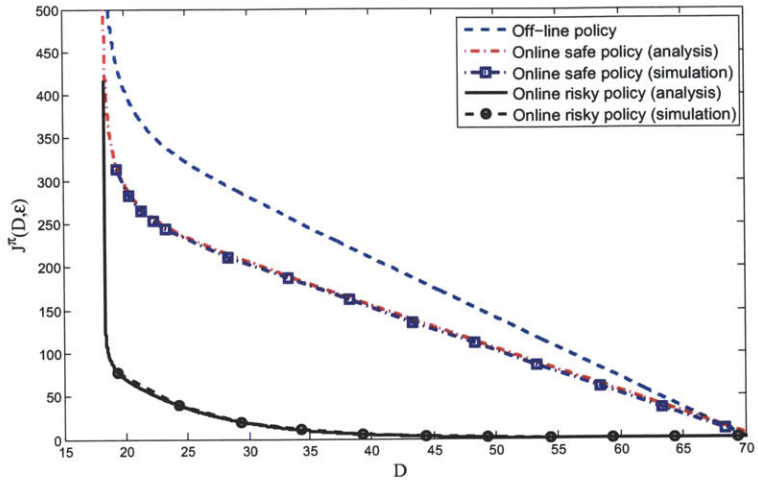

Figure 4-4: Expected cost (units of time) of the presented control policies as a function of the initial buffer size for interruption probability $\epsilon = 10^{-3}$. The analytical bounds are given by Theorems 4.2.2, 4.2.3 and 4.2.5.

online safe and off-line policies. For example, the risky policy allows us to decrease the initial buffer size from 70 to 20 with an average of $70 \times 0.15 \approx 10$ extra packets from the costly server. The expected cost in terms of the number packets received from the costly server is 43 and 61 for the online safe and off-line policy, respectively.

Moreover, note that it is merely the existence of the costly server as a backup that allows us to improve the user's quality of experience without actually using too many packets from the costly server. For example, observe that the risky policy satisfies QoE metrics of $D = 35$ and $\epsilon = 10^{-3}$, by only using on average about one extra packet from the costly server. However, without the costly server, in order to decrease the initial buffer size from 70 to 35, the interruption probability has to increase from 10^{-3} to about 0.03 (see Corollary 2.2.2).

4.3 Dynamic Programming Approach

In this section, we present a characterization of the optimal association policy in terms of the Hamilton-Jacobi-Bellman (HJB) equation. Note that because of the probabilistic constraint over the space of sample paths of the buffer size, the optimal policy is not necessarily Markov with respect to the buffer size as the state of the

system. We take a similar approach as in [5] where by expanding the state space, a Bellman equation is provided as the optimality condition of an MDP with probabilistic constraint. In particular, consider the pair (Q_t, p_t) as the state variable, where Q_t denotes the buffer size and p_t represents the desired level of interruption probability given the information at time t. Note that p_t is a martingale by definition [44]. The evolution of Q_t is governed by the following stochastic differential equation

$$dQ_t = -dt + dN^u, \quad Q_0 = D, \tag{4.14}$$

where N^u is a Poisson counter with rate $R_u = R_0 + u_t \cdot R_c$. For any $(D, \epsilon) \in \mathcal{R}$ and any optimal policy π, the constraint $p^\pi(D) \leq \epsilon$ is active. Hence, we consider the sample paths of p_t such that $p_0 = \epsilon$. Moreover, we have $\mathbf{E}[p_t] = \epsilon$ for all t, where the expectation is with respect to the Poisson jumps. Let $dp_t = \hat{p}_t - p_t$ be the change in state p, if a Poisson jump occurs in an infinitesimal interval of length dt. Also, let $dp_t = dp_0$ be the change in state p if no jump occurs. Therefore,

$$0 = \mathbf{E}[dp_t] = R_u dt(\hat{p}_t - p_t) + (1 - R_u dt)dp_0.$$

By solving the above equation for dp_0, we obtain the evolution of p as a function of the control process $\hat{p}_t$ and u_t:

$$dp_t = (p_t - \hat{p}_t)(R_u dt - dN^u), \quad p_0 = \epsilon. \tag{4.15}$$

Similarly to the arguments of Theorem 2 of [5], by principle of optimality we can write the following dynamic programming equation

$$V(Q,p) = \min_{u \in \{0,1\}, \hat{p} \in [0,1]} \big\{ u dt + \mathbf{E}[V(Q + dQ, p + dp)] \big\}. \tag{4.16}$$

If V is continuously differentiable, by Itō's Lemma for jump processes, we have

$$V(Q + dQ, p + dp) - V(Q,p) = \frac{\partial V}{\partial Q}(-dt) + \frac{\partial V}{\partial p} \cdot (p - \hat{p}) R_u dt + \big(V(Q+1, \hat{p}) - V(Q,p)\big) dN^u,$$

which implies the following HJB equation after dividing (4.16) by dt and taking the limit as t goes to zero:

$$\frac{\partial V(Q,p)}{\partial Q} = \min_{u\in\{0,1\},\hat{p}\in[0,1]} \{u + \frac{\partial V}{\partial p}\cdot(p-\hat{p})R_u + R_u\big(V(Q+1,\hat{p}) - V(Q,p)\big)\}$$

The optimal policy π is obtained by characterizing the optimal solution of the partial differential equation in (4.17) together with the boundary condition $V(Q,1) = 0$. Since such equations are in general difficult to solve analytically, we use the *guess and check* approach, where we propose a candidate for the value function and verify that it nearly satisfies the HJB equation almost everywhere. Moreover, we show that the trajectories of (Q_t, p_t) steered by the optimal actions $(u^*, \hat{p}^*)$ lie in a one-dimensional invariant manifold, leading to the risky policy defined in Definition 4.2.2.

For any $(Q,p) \in \mathcal{R}$ define

$$T(Q,p) = \begin{cases} \frac{1}{\alpha_1-\alpha_0}\big[\log\left(\frac{\theta}{p}\right) - \alpha_0 Q\big], & \text{if } Q \geq \frac{1}{a_1}\log\left(\frac{\theta}{p}\right), \\ \frac{1}{\alpha_1}\log\left(\frac{p+\theta(1-e^{-\alpha_1 Q})-1}{p-e^{-\alpha_1 Q}}\right), & \text{otherwise,} \end{cases} \tag{4.17}$$

where $\theta = \frac{\alpha_1}{\alpha_0}$. The candidate solution for HJB equation (4.17) is given by

$$\bar{V}(Q,p) = \frac{1}{\alpha_0(1-\frac{\alpha_0}{2})(R_1-1)} e^{-\alpha_0(Q-T(Q,p))}, \tag{4.18}$$

when $Q \geq \frac{1}{a_1}\log\left(\frac{\theta}{p}\right)$, and

$$\bar{V}(Q,p) = \frac{p+\theta(1-e^{-\alpha_1 Q})-1}{(R_1-1)(\theta-1)}\big(T(Q,p) + \frac{\beta}{\alpha_1}\big) - \frac{Q}{R_1-1}, \tag{4.19}$$

when $Q < \frac{1}{a_1}\log\left(\frac{\theta}{p}\right)$. Note that the candidate solution is derived from the structure of the expected cost of the risky policy (cf. Theorem 4.2.5). We may verify that $\bar{V}$ satisfies the HJB equation (4.17) for all (Q,p) such that $Q \geq \frac{1}{a_1}\log\left(\frac{\theta}{p}\right)$ or $Q \leq \frac{1}{a_1}\log\left(\frac{\theta}{p}\right) - 1$, but for other (Q,p) the HJB equation is only approximately satisfied. This is due to the discontinuity of the queue-length process which does not allow us to exactly match the expected cost starting from below the threshold with the one

starting from above the threshold. Therefore, due to approximate characterization of the cost of the risky policy, we may not prove or disprove optimality of this policy.

In order to provide further intuition behind this formulation and the risky policy, we consider the sample paths of the state process. We expect that for the initial condition $(Q_0, p_0) = (D, \epsilon)$, the trajectory of (Q_t, p_t) steered by the optimal actions $(u_t^*, \hat{p}_t^*)$ is limited to a one-dimensional invariant manifold $\mathcal{M}(D, \epsilon)$, where

$$\mathcal{M}(D,\epsilon) = \Big\{(Q,p) : p = \theta e^{-\alpha_0 Q - (\alpha_1 - \alpha_0)T(D,\epsilon)} \cdot \mathbb{I}_{\{Q \geq T(D,\epsilon)\}} + \frac{(\theta-1)e^{-\alpha_1 T(D,\epsilon)} + e^{-\alpha_1 Q}(1 - \theta e^{-\alpha_1 T(D,\epsilon)})}{1 - e^{-\alpha_1 T(D,\epsilon)}} \mathbb{I}_{\{Q < T(D,\epsilon)\}}\Big\}, \quad (4.20)$$

where $T(D, \epsilon)$ is given by (4.17). Figure 4-5 plots these one-dimensional manifolds for different initial conditions (D, ϵ). The system state crossing from the region with $u^* = 0$ to the region with $u^* = 1$, is equivalent to the queue-length crossing a threshold. Since the system state is limited to a one-dimensional space, this threshold remains constant over time.

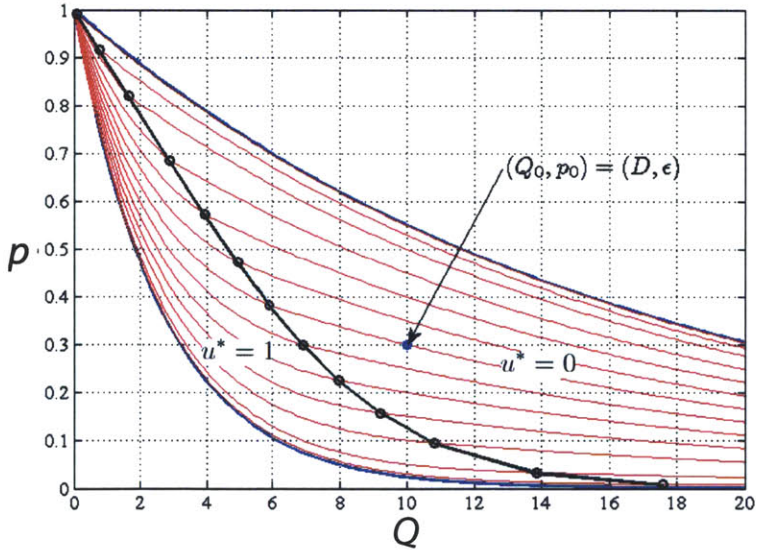

Figure 4-5: Trajectory of the optimal policy lies on a one-dimensional manifold.

In the following, we use a fluid model to provide an *exact* characterization of the optimal control policy using appropriate HJB equation.

4.4 Optimal Association Policy for a Fluid Model

Thus far, we concentrated on design and analysis of various network association policies in an uncertain environment, where network uncertainties are modeled using a Poisson arrival process. We provided closed-form approximations of the cost of different policies. However, an exact analytical solution is required to prove optimality of the risky policy. This is particularly challenging since an exact distribution of threshold over-shoots is desired due to discontinuous nature of the Poisson process. In this part, we exploit a second-order approximation of the Poisson process [51] and model the receiver's buffer size using a controlled Brownian motion with drift.

Consider the system model as in Figure 4-1, with following queue-length dynamics at the receiver:

$$dQ_t = (R_{u_t} - 1)dt + dW_t, \quad Q_0 = D, \tag{4.21}$$

where W_t is the Wiener process, $u_t \in \{0, 1\}$ is the receiver's decision at time t on using the free or costly server. As in the preceding part, we assume that the media file size F is infinite and $R_1 > R_0 > 1$.

Define the control policy (network association policy) as in Definition 4.1.1. The goal is to find a feasible policy that minimizes the usage cost defined in (4.3) such that the interruption probability $p^\pi(D)$ defined in (4.2) is at most ϵ. As in the previous part, the set of feasible policies and the value function is given by Definition 4.1.2 and (4.4), respectively. The following lemma is the counterpart of Corollary 2.2.2 for the fluid model.

Lemma 4.4.1. *Let $\pi_i \equiv i$ denote a degenerate policy, where $i \in \{0, 1\}$. The interruption probability for such policy is given by*

$$p^{\pi_i}(D) = e^{-\theta_i D}, \quad \textit{for all } D \geq 0, \tag{4.22}$$

where $\theta_i = 2(R_i - 1)$, for $i \in \{0, 1\}$.

Proof. See Appendix 4.6 at the end of this chapter. □

First, we provide a characterization of the optimal policy via Hamilton-Jacobi-Bellman equation. As in Section 4.3, we expand the state variables to (Q, p), where Q is the queue-length with dynamics given by 4.21, and p is the desired interruption probability. Using Martingale representation theorem [44], we may write the dynamics of p as follows:

$$dp = \hat{p}_t \; dW_t, \quad p_0 = \epsilon, \tag{4.23}$$

where $\hat{p}_t$ is a predictable process which is adapted with respect to natural filtration of the history process. Using the principal of optimality, we may write the following dynamic programming equation:

$$V(Q,p) = \min_{(u,\hat{p})\in\{0,1\}\times\mathbb{R}} \big\{ udt + \mathbf{E}[V(Q+dQ, p+dp)] \big\}, \tag{4.24}$$

where $(u, \hat{p})$ are the control actions. For a twice differentiable function V, we may exploit Itō's Lemma to get

$$\begin{aligned} dV(Q,p) &= \frac{\partial V}{\partial Q}dQ + \frac{\partial V}{\partial p}dp + \frac{1}{2}\frac{\partial^2 V}{\partial Q^2}(dQ)^2 + \frac{1}{2}\frac{\partial^2 V}{\partial p^2}(dp)^2 + \frac{\partial^2 V}{\partial Q \partial p}dQdp \\ &\overset{(a)}{=} \frac{\partial V}{\partial Q}((R_u - 1)dt + dW) + \frac{\partial V}{\partial p}(\hat{p}\; dW) \\ &\quad + \frac{1}{2}\frac{\partial^2 V}{\partial Q^2}dt + \frac{1}{2}\frac{\partial^2 V}{\partial p^2}(\hat{p})^2 dt + \frac{\partial^2 V}{\partial Q \partial p}\hat{p}dt, \end{aligned}$$

where (a) follows from state dynamics in (4.21) and (4.23). Replacing the above equation back in (4.24), and taking the expectation with respect to dW, and limit as dt tends to zero, we obtain the following HJB equation

$$0 = \min_{(u,\hat{p})\in\{0,1\}\times\mathbb{R}} \left\{ u + \frac{\partial V}{\partial Q}(R_u - 1) + \frac{1}{2}\frac{\partial^2 V}{\partial Q^2} + \frac{1}{2}\frac{\partial^2 V}{\partial p^2}(\hat{p})^2 + \frac{\partial^2 V}{\partial Q \partial p}\hat{p} \right\}. \tag{4.25}$$

Note that we require the following boundary conditions for the value function:

$$V(Q,1) = V(0,p) = 0, \quad \text{for all } Q \geq 0, 0 \leq p \leq 1 \tag{4.26}$$

Providing an analytical solution for the partial differential equation in (4.25) is

often challenging or impossible. However, we may take a guess and check approach and use a threshold policy as the basis of our guess. Note that we need to verify the HJB equation for the set of state variables that reachable by a feasible policy. In particular, in light of Lemma 4.4.1, it is clear that for $p \leq e^{-\theta_1 Q}$, there is no feasible policy and the value function $V(Q,p) = \infty$. Moreover, for all $p \geq e^{-\theta_0 Q}$, observe that the degenerate policy $\pi_0 \equiv 0$ is optimal, $V(Q,p) = 0$ which also satisfies the HJB equation and the boundary conditions. Therefore, we focus on the non-degenerate region

$$\mathcal{R} = \Big\{(Q,p) : e^{-\theta_1 Q} < p < e^{-\theta_0 Q}, Q \geq 0\Big\}. \tag{4.27}$$

Figure 4-2 illustrates a conceptual example of this non-degenerate region. In the following, we first define a threshold policy similar to the risky policy of Definition 4.2.2, and present a closed-form characterization of its cost function. Then, we show that for a proper choice of the threshold the associated cost function satisfies the HJB equation in (4.25) and the optimal solution of the minimization problem in (4.25) coincides with the threshold policy. Hence, we establish the optimality of the proposed policy.

Theorem 4.4.1. *Let π^T be the threshold policy as in Definition 4.2.2, parameterized with threshold value T. Also, let the queue-length dynamics be governed by (4.21). Then, the interruption probability for this policy is given by*

$$p^T(D) = \begin{cases} e^{-\theta_1 D} + p(T)(1 - \frac{\theta_0}{\theta_1})\big(1 - e^{-\theta_1 D}\big), & 0 \leq D \leq T \\ p(T)e^{-\theta_0(D-T)}, & D \geq T, \end{cases} \tag{4.28}$$

where

$$p(T) = \frac{e^{-\theta_1 T}}{\frac{\theta_0}{\theta_1} + (1 - \frac{\theta_0}{\theta_1})e^{-\theta_1 T}},$$

and $\theta_i = 2(R_i - 1)$, for $i \in \{0,1\}$.

Proof. See Appendix 4.6 at the end of this chapter. □

Corollary 4.4.1. *Let π^T be the threshold policy as in Definition 4.2.2. Then, the policy π^T is (D,ϵ)-feasible (cf. Definition 4.1.2) for the following choices of the threshold*

T:

1. *For all* $\epsilon \geq e^{-\theta_0 D}$, *let* $T = 0$.

2. *For all* $e^{-\theta_1 D} \leq \epsilon \leq e^{-\theta_0 D}$, *let* $T = T(D, \epsilon)$ *be the unique solution of* $p^T(D) = \epsilon$, *where* $p^T(D)$ *is given by (4.28).*

3. *For all other* ϵ, *there exists no such* T.

Proof. The proof directly follows from the characterization of the interruption probability in Theorem 4.4.1, noting the fact that $p^T(D) \in [e^{-\theta_1 D}, e^{-\theta_0 D}]$ is monotonically decreasing in T. □

Next, we provide a exact characterization of the expected cost of the threshold policy π^T for a given threshold T. This allows us to obtain a proper candidate solution for the HJB equation.

Theorem 4.4.2. *Let* π^T *be the threshold policy as in Definition 4.2.2. Define* $J^T(D)$ *as the expected cost associated with policy* π^T *given the initial condition* D *for queue-length dynamics (4.21) and threshold* T. *The* cost-to-go *function* $J^T(D)$ *is given by*

$$J^T(D) = \begin{cases} e^{-\theta_0(D-T)} J(T), & D \geq T \\ (J(T) + \frac{2}{\theta_1} T) \frac{1-e^{-\theta_1 D}}{1-e^{-\theta_1 T}} - \frac{2}{\theta_1} D, & D \leq T, \end{cases} \tag{4.29}$$

where

$$J(T) = \frac{\frac{2}{\theta_1^2}\left[1 - (1+\theta_1 T)e^{-\theta_1 T}\right]}{\frac{\theta_0}{\theta_1} + (1 - \frac{\theta_0}{\theta_1})e^{-\theta_1 T}}. \tag{4.30}$$

Proof. See Appendix 4.6 at the end of this chapter. □

The following theorem provide a candidate for value function and verify the optimality condition given by HJB equation in (4.25).

Theorem 4.4.3. *For all* $(Q, p) \in \mathcal{R}$, *define*

$$V(Q, p) = J^{T(Q,p)}(Q), \tag{4.31}$$

where $J^T(\cdot)$ is defined in (4.29), $\mathcal{R}$ is defined in (4.27), and $T(Q,p)$ is the unique solution of

$$p^T(Q) = p.$$

Then, the HJB equation (4.25) and boundary condition (4.26) hold for all $(Q,p) \in \mathcal{R}$.

Proof. See Appendix 4.6 at the end of this chapter. □

Theorem 4.4.3 verifies that the value function $V(Q,p)$ given by (4.31) is indeed the optimal cost function defined in (4.4). Furthermore, we can conclude that the policy $\pi^*(Q,p)$ achieving the minimum in the HJB equation (4.25) is optimal. In general, the optimal policy depends on both state variables (Q,p) and is not Markov with respect to Q. In the following, we show that the state trajectory steered by the optimal policy is limited to a one-dimensional manifold and the threshold policy π^T is optimal for all $(Q,p) \in \mathcal{R}$. Recall the policy π^T boils down to the optimal policy $\pi_0 \equiv 0$ for all other admissible states, by using threshold value $T = 0$.

Theorem 4.4.4. *Let $\pi^*(Q,p)$ attain the minimum in the HJB equation (4.25) for any $(Q,p) \in \mathcal{R}$. Let (Q_t^*, p_t^*) denote the state trajectory given the initial condition (D,ϵ), under the control trajectory $(u_t^*, \hat{p}_t^*) = \pi^*(Q_t^*, p_t^*)$. Then, the state trajectory is limited to a one-dimensional invariant manifold $\mathcal{M}(D,\epsilon)$, where*

$$\mathcal{M}(D,\epsilon) = \left\{(Q,p) : p = p^{T(D,\epsilon)}(Q)\right\}, \tag{4.32}$$

where $T(D,\epsilon)$ is the solution of $p^T(D) = \epsilon$, and $p^T(\cdot)$ is defined in (4.28). Moreover, the optimal policy $\pi^*(Q,p)$ coincides with the threshold policy $\pi^{T(D,\epsilon)}(Q)$.

Proof. See Appendix 4.6 at the end of this chapter. □

Corollary 4.4.2. *The optimal policy $\pi^*(Q,p)$ is Markovian with respect to the queue-length process Q_t conditioned on the initial condition $(Q_0, p_0) = (D,\epsilon)$.*

Figure 4-5 illustrates a conceptual figure of the one-dimensional invariant manifolds for different initial conditions, which contain the optimal state trajectory and lead to optimality of the threshold policy.

4.5 Appendix to Chapter 4 - Analysis of the Control Policies for the Poisson Arrival Model

Proof of Theorem 4.2.2. By Definition 4.1.2, we need to show that $p^{\pi}(D) \leq \epsilon$. By a union bound on the interruption probability, it is sufficient to verify

$$\mathbf{Pr}\Big(\min_{0\leq t\leq t_s} Q_t \leq 0 \big| Q_0 = D\Big) + \mathbf{Pr}\Big(\min_{t>t_s} Q_t \leq 0 \big| Q_0 = D\Big) \leq \epsilon. \tag{4.33}$$

In the interval $[0, t_s]$, Q_t behaves as in a single-server system with rate R_1. Hence, by Corollary 2.2.2 we get

$$\mathbf{Pr}\Big(\min_{0\leq t\leq t_s} Q_t \leq 0 \big| Q_0 = D\Big) \leq e^{-\alpha_1 D}. \tag{4.34}$$

For the second term in (4.33), we have

$$\begin{aligned}
\mathbf{Pr}\Big(\min_{t>t_s} Q_t \leq 0 \big| Q_0 = D\Big) &= \sum_{q=D-t_s}^{\infty} \mathbf{Pr}\Big(\min_{t>t_s} Q_t \leq 0 \big| Q_{t_s} = q\Big)\mathbf{Pr}(Q_{t_s} = q) \\
&\overset{(a)}{\leq} \sum_{q=D-t_s}^{\infty} e^{-\alpha_0 q}\mathbf{Pr}(Q_{t_s} = q) \\
&= \sum_{k=0}^{\infty} e^{-\alpha_0(D+k-t_s)}\mathbf{Pr}(N_{t_s} + N^c_{t_s} = k) \\
&\overset{(b)}{=} \sum_{k=0}^{\infty} e^{-\alpha_0(D+k-t_s)}\frac{e^{-R_1 t_s}(R_1 t_s)^k}{k!} \\
&= e^{-\alpha_0(D-t_s)+R_1 t_s(e^{-\alpha_0}-1)} \sum_{k=0}^{\infty} \frac{e^{-R_1 t_s e^{-\alpha_0}}(R_1 t_s e^{-\alpha_0})^k}{k!} \\
&= \exp\Big(-\alpha_0(D-t_s) + R_1 t_s(e^{-\alpha_0}-1)\Big)\cdot 1 \\
&\overset{(c)}{=} \exp\Big(-\alpha_0(D-t_s) + R_1 t_s(-\frac{\alpha_0}{R_0})\Big) \\
&\overset{(d)}{\leq} \epsilon - e^{-\alpha_1 D},
\end{aligned}$$

where (a) follows from Corollary 2.2.2 and the fact that $u_t = 0$, for $t \geq t_s$. (b) is

true because $N_{t_s} + N^c_{t_s}$ is a Poisson random variable with mean $R_1 t_s$. (c) holds since $\alpha_0 = I(R_0)$ is the root of $\gamma(r) = r + R_0(e^{-r} - 1)$. Finally, (d) follows from the hypothesis of the theorem.

By combining the above bounds, we may verify (4.33) which in turns proves feasibility of the proposed control policy. ■

Proof of Theorem 4.2.3. Similarly to the proof of Theorem 4.2.2, we need to show that the total probability of interruption before and after crossing the threshold S is bounded from above by ϵ. Observe that for any realization of τ_S the bound in (4.34) still holds. Further, since the costly server is not used after crossing the threshold and $Q_{\tau_S} \geq S$, Corollary 2.2.2 implies

$$\mathbf{Pr}\Big(\min_{t>\tau_S} Q_t \leq 0 \big| Q_0 = D\Big) \leq e^{-\alpha_0 S} \leq \epsilon - e^{-\alpha_1 D}, \tag{4.35}$$

where the second inequality follows from (4.9). Finally, combining (4.34) and (4.35) gives $p^{\pi^S}(D) \leq \epsilon$, which is the desired feasibility result.

For the second part, first observe that $J^{\pi^S}(D, \epsilon) = \mathbf{E}[\tau_S]$. In order to cross a threshold $S \geq S^*$, the threshold S^* must be crossed earlier, because $Q_0 = D \leq S^*$. Hence, τ_S stochastically dominates τ^*_S, implying

$$J^{\pi^S}(D, \epsilon) = \mathbf{E}[\tau_S] \geq \mathbf{E}[\tau_{S^*}] = J^{\pi^{S^*}}(D, \epsilon), \quad \text{for all } S \geq S^*.$$

It only remains to compute $\mathbf{E}[\tau_{S^*}]$. It follows from Wald's identity or Doob's optional stopping theorem [44] that

$$D + (R_1 - 1)\mathbf{E}[\tau_{S^*}] = \mathbf{E}[Q_{\tau_{S^*}}] = S^* + \xi, \tag{4.36}$$

where $\xi \in [0, 1)$ because the jumps of a Poisson process are of units size, and hence the overshoot size when crossing a threshold is bounded by one, i.e., $S^* \leq Q_{\tau_{S^*}} < S^* + 1$. Rearranging the terms in (4.36) and plugging the value of S^* from (4.9) immediately gives the result. ■

Lemma 4.5.1. *Let Q_t be the buffer size of a single-server system with arrival rate $R > 1$. Let the initial buffer size be D and for any $T \geq D > 0$ define the following stopping times*

$$\tau_T = \inf\{t > 0 : Q_t \geq T\}, \quad \tau_e = \inf\{t \geq 0 : Q_t \leq 0\}. \tag{4.37}$$

Then

$$\mathbf{Pr}(\tau_e > \tau_T) = \frac{1 - e^{-I(R)D}}{1 - \mathbf{E}[e^{-I(R)Q_{\tau_T}} | \tau_e > \tau_T]}, \tag{4.38}$$

where $I(R)$ is defined in (2.5).

Proof. Let $Y(t) = e^{-I(R)Q_t}$. We may verify that $Y(t)$ is a martingale and uniformly integrable. Also, define the stopping time $\tau = \min\{\tau_T, \tau_e\}$. Since $R > 1$, we have $\mathbf{Pr}(\tau \geq t) \leq \mathbf{Pr}(0 < Q_t < T) \to 0$, as $t \to \infty$. Hence, $\tau < \infty$ almost surely. Therefore, we can employ Doob's optional stopping theorem [44] to write

$$\begin{aligned} e^{-I(R)D} &= \mathbf{E}[Y(0)] = \mathbf{E}[Y(\tau)] \\ &= \mathbf{Pr}(\tau_e \leq \tau_T) \cdot 1 \\ &\quad + \mathbf{Pr}(\tau_e > \tau_T)\mathbf{E}[e^{-I(R)Q_{\tau_T}} | \tau_e > \tau_T]. \end{aligned}$$

The claim immediately follows from the above relation after rearranging the terms.

□

Proof of Theorem 4.2.4. Let us first characterize the interruption probability of the policy π^T when the initial buffer size is $D = T$. In this case, by definition of π^T the behavior of Q_t is initially the same as a single-server system with rate R_1 until

the threshold T is crossed. Hence, by Lemma 4.5.1 we have

$$\begin{aligned} p^{\pi^T}(T) &= \mathbf{Pr}\Big(\min_{t\geq 0} Q_t \leq 0 \big| Q_0 = T\Big) \\ &= \mathbf{Pr}(\tau_e < \tau_T)\cdot 1 \\ &\quad + \mathbf{Pr}(\tau_T < \tau_e)\mathbf{Pr}\Big(\min_{t\geq \tau_T} Q_t \leq 0 \big| \tau_T < \tau_e, Q_0 = T\Big) \\ &= \frac{e^{-\alpha_1 T} - \mathbf{E}[e^{-\alpha_1 Q_{\tau_T}}|\tau_e > \tau_T]}{1 - \mathbf{E}[e^{-\alpha_1 Q_{\tau_T}}|\tau_e > \tau_T]} \\ &\quad + \frac{\big(1 - e^{-\alpha_1 T}\big)\mathbf{Pr}\Big(\min_{t\geq \tau_T} Q_t \leq 0 \big| \tau_T < \tau_e, Q_0 = T\Big)}{1 - \mathbf{E}[e^{-\alpha_1 Q_{\tau_T}}|\tau_e > \tau_T]}. \end{aligned} \tag{4.39}$$

Further,

$$\begin{aligned} &\mathbf{Pr}\Big(\min_{t\geq \tau_T} Q_t \leq 0 \big| \tau_T < \tau_e, Q_0 = T\Big) = \int_T^{T+1} \mathbf{Pr}\Big(\min_{t\geq \tau_T} Q_t \leq 0 \big| Q_{\tau_T}\Big) d\mu(Q_{\tau_T}) \\ &\overset{(a)}{=} \int_T^{T+1} \mathbf{Pr}\Big(\min_{t\geq 0} Q_t \leq 0 \big| Q_0\Big) d\mu(Q_0) \\ &\overset{(b)}{=} \int_T^{T+1} \mathbf{Pr}\Big(\min_{t\geq 0} Q_t \leq 0 \big| \min_{t\geq 0} Q_t \leq T, Q_0\Big)\mathbf{Pr}\Big(\min_{t\geq 0} Q_t \leq T | Q_0\Big) d\mu(Q_0) \\ &\overset{(c)}{=} \int_T^{T+1} p^{\pi^T}(T) e^{-\alpha_0(Q_0 - T)} d\mu(Q_0) \\ &= \mathbf{E}[e^{-\alpha_0(Q_{\tau_T} - T)}|\tau_T < \tau_e] p^{\pi^T}(T), \end{aligned} \tag{4.40}$$

where μ denotes the conditional distribution of Q_{τ_T} given $\tau_T < \tau_e$. Note that $Q_{\tau_T} \in [T, T+1]$ because the size of the overshoot is bounded by one. Further, (a) follows from stationarity of the arrival processes and the control policy, (b) holds because a necessary condition for the interruption event is to cross the threshold T when starting from a point $Q_0 \geq T$. Finally (c) follows from Corollary 2.2.2 and the definition of the risky policy. The relations (4.39) and (4.40) together result in

$$p^{\pi^T}(T) = \frac{e^{-\alpha_1 T}\big(1 - \mathbf{E}_\mu[e^{-\alpha_1(Q_{\tau_T} - T)}]\big)}{1 - \mathbf{E}_\mu[e^{-\alpha_0(Q_{\tau_T} - T)}] + \kappa}, \tag{4.41}$$

where $\kappa = \mathbf{E}_\mu[e^{-\alpha_0 Q_{\tau_T} - (\alpha_1 - \alpha_0)T}] - \mathbf{E}_\mu[e^{-\alpha_1 Q_{\tau_T}}] \geq 0$. Therefore, using the fact that

$$1 - x \leq e^{-x} \leq 1 - x + \frac{x^2}{2}, \quad \text{for all } x \geq 0, \tag{4.42}$$

we can provide the following bound

$$\begin{aligned} p^{\pi^T}(T) &\leq \frac{e^{-\alpha_1 T}\big(\alpha_1 \mathbf{E}_\mu[Q_{\tau_T} - T]\big)}{\alpha_0 \mathbf{E}_\mu[Q_{\tau_T} - T]\Big(1 - \frac{\alpha_0}{2} \cdot \frac{\mathbf{E}_\mu[(Q_{\tau_T} - T)^2]}{\mathbf{E}_\mu[Q_{\tau_T} - T]}\Big)} \\ &\leq \frac{\alpha_1}{\alpha_0(1 - \frac{\alpha_0}{2})} e^{-\alpha_1 T} \\ &= \beta e^{-\alpha_1 T}, \end{aligned} \tag{4.43}$$

where the last inequality holds, since $0 \leq Q_{\tau_{\bar{D}}} - \bar{D} \leq 1$.

Now we prove feasibility of the risky policy π^{T^*} when $D > \bar{D}$. Observe that by (4.11), $D > T^*$, hence the behavior of the buffer size Q_t is the same as the one in a single-server system with rate R_0 until the threshold T^* is crossed. Thus

$$\begin{aligned} p^{\pi^{T^*}}(D) &= \mathbf{Pr}\Big(\min_{t \geq 0} Q_t \leq 0 \big| Q_0 = D\Big) \\ &= \mathbf{Pr}\Big(\min_{t \geq 0} Q_t \leq 0 \big| \min_{t \geq 0} Q_t \leq T^*, Q_0 = D\Big)\mathbf{Pr}\Big(\min_{t \geq 0} Q_t \leq T^* | Q_0 = D\Big) \\ &= p^{\pi^{T^*}}(T^*)e^{-\alpha_0(D - T^*)} \\ &\leq \beta e^{-(\alpha_1 - \alpha_0)T^* - \alpha_0 D} = \epsilon, \end{aligned}$$

where the inequality follows from (4.43), and the last equality holds by (4.11).

Next we verify the feasibility of the policy π^{T^*} for $D \leq \bar{D}$. In this case, $D \leq T^*$ and by definition of the risky policy the system behaves as a single-server system with arrival rate R_1 until the threshold T^* is crossed or the buffer size hits zero

(interruption). Hence, we can bound the interruption probability as follows

$$\begin{aligned} p^{\pi^{T^*}}(D) &= \mathbf{Pr}(\tau_e < \tau_{T^*}) \cdot 1 + \mathbf{Pr}(\tau_{T^*} < \tau_e)\mathbf{Pr}\Big(\min_{t \geq \tau_{T^*}} Q_t \leq 0 \Big| \tau_{T^*} < \tau_e, Q_0 = D\Big) \\ &\overset{(a)}{=} 1 - \mathbf{Pr}(\tau_{T^*} < \tau_e)\Big(1 - \mathbf{E}_\mu[e^{-\alpha_0(Q_{\tau_{T^*}} - T^*)}]p^{\pi^{T^*}}(T^*)\Big) \\ &\overset{(b)}{\leq} \frac{(\beta-1)(1-e^{-\alpha_1 D})}{1 - \mathbf{E}_\mu[e^{-\alpha_1 Q_{\tau_{T^*}}}]} + 1 - \beta(1 - e^{-\alpha_1 D}) \\ &\overset{(c)}{\leq} \frac{(\beta-1)(1-e^{-\alpha_1 D})}{1 - e^{-\alpha_1 T^*}} + 1 - \beta(1 - e^{-\alpha_1 D}) \overset{(d)}{=} \epsilon, \end{aligned}$$

where (a) follows from (4.40), (b) can be verified after some manipulations by combining the result of Lemma 4.5.1 and (4.41), and (c) holds since $\beta \geq 1$ and $Q_{\tau_{T^*}} \geq T^*$. Finally, (d) immediately follows from plugging in the definition of T^* from (4.11).

Therefore, the risky policy π^{T^*} is feasible by Definition 4.1.2. Observe that the buffer size under any policy π^T of the form (4.10) with $T \geq T^*$ stochastically dominates that of policy π^{T^*}, because π^T switches to the costly server earlier, and stays in that state longer. Hence, π^T is feasible for all $T \geq T^*$. ■

Proof of Theorem 4.2.5. Similarly to the proof of Theorem 4.2.4, we first consider the risky policy π^T with the initial buffer size T. By definition of π^T, the costly server is used until the threshold T is crossed. Thus the expected cost of this policy is bounded by the expected time until crossing the threshold plus the expected cost given that the threshold is crossed, i.e.,

$$J^{\pi^T}(T,\epsilon) \leq \frac{\mathbf{E}[Q_{\tau_T}] - T}{R_1 - 1} + E[e^{-\alpha_0(Q_{\tau_T} - T)}]J^{\pi^T}(T,\epsilon),$$

where τ_T is defined in (4.37). The above relation implies

$$\begin{aligned} J^{\pi^T}(T,\epsilon) &\leq \frac{1}{R_1 - 1} \cdot \frac{\mathbf{E}[Q_{\tau_T} - T]}{1 - E[e^{-\alpha_0(Q_{\tau_T} - T)}]} \\ &\leq \frac{1}{R_1 - 1} \cdot \frac{1}{\alpha_0\Big(1 - \frac{\alpha_0}{2} \cdot \frac{\mathbf{E}_\mu[(Q_{\tau_T} - T)^2]}{\mathbf{E}_\mu[Q_{\tau_T} - T]}\Big)} \\ &\leq \frac{1}{\alpha_0(R_1 - 1)(1 - \frac{\alpha_0}{2})} = \frac{\beta}{\alpha_1(R_1 - 1)}, \end{aligned} \tag{4.44}$$

where the second inequality follows from the fact in (4.42). Now for any $D \geq \bar{D}$ we can write

$$\begin{aligned} J^{\pi^{T^*}}(D,\epsilon) &= \mathbf{Pr}\Big(\min_{t\geq 0} Q_t \leq T^* \big| Q_0 = D\Big) J^{\pi^{T^*}}(T^*,\epsilon) \\ &= e^{-a_0(D-T^*)} J^{\pi^{T^*}}(T^*,\epsilon) \end{aligned}$$

where the inequality holds by Corollary 2.2.2. Combining this with (4.44) gives the result in (4.12).

If $D \leq \bar{D}$, the risky policy uses the costly server until the threshold T^* is crossed at τ_{T^*} or the interruption event (τ_e), whichever happens first. Afterwards, no extra cost is incurred if an interruption has occurred. Otherwise, by (4.44) an extra cost of at most $\frac{\beta}{\alpha_1(R_1-1)}$ is incurred, i.e.,

$$J^{\pi^{T^*}}(D,\epsilon) \leq \mathbf{E}\big[\min\{\tau_e, \tau_{T^*}\}\big] + \mathbf{Pr}(\tau_{T^*} < \tau_e)\frac{\beta}{\alpha_1(R_1-1)}.$$

By Doob's optional stopping theorem applied to the martingale $Z_t = Q_t - (R_1-1)t$, we obtain

$$D = \mathbf{Pr}(\tau_{T^*} < \tau_e)\mathbf{E}[Q_{\tau_{T^*}} | \tau_{T^*} < \tau_e] - (R_1-1)\mathbf{E}\big[\min\{\tau_e, \tau_{T^*}\}\big],$$

which implies

$$\mathbf{E}\big[\min\{\tau_e, \tau_{T^*}\}\big] \leq \frac{\mathbf{Pr}(\tau_{T^*} < \tau_e)(T^*+1) - D}{R_1 - 1}$$

By combining the preceding relations we conclude that

$$J^{\pi^{T^*}}(D,\epsilon) \leq \frac{\mathbf{Pr}(\tau_{T^*} < \tau_e)}{R_1 - 1}\Big(T^* + 1 + \frac{\beta}{\alpha_1}\Big) - \frac{D}{R_1 - 1},$$

which immediately implies (4.13) by employing Lemma 4.5.1.

■

4.6 Appendix to Chapter 4 - Analysis of the Threshold Policy and HJB equation for the Fluid Approximation Model

Proof of Lemma 4.4.1. There are multiple approaches to prove the claim. We prove a more general case using Doob's optional stopping theorem that would be useful in the later arguments. We only consider $i = 0$; the other case is the same.

Let $Y_t = e^{-\theta_0 Q_t}$. It is straightforward to show that Y_t is a martingale with respect to W_t. Now consider the boundary crossing problem, where we are interested in the probability of hitting zero before a boundary $b > D$. Let τ denote the hitting time of either boundaries. For any $n > 0$, we may apply Doob's optional stopping theorem [44] to the stopped martingale $Y_{\tau \wedge n}$ to write

$$\mathbf{E}[Y_{\tau \wedge b}] = \mathbf{E}[e^{-\theta_0 Q_{\tau \wedge n}}] = e^{-\theta_0 D}, \quad \text{for all } n.$$

Now, we take the limit as $n \to \infty$ and exploit the dominant convergence theorem to establish:

$$\mathbf{E}[Y_\tau] = \mathbf{E}[e^{-\theta_0 Q_\tau}] = \lim_{n \to \infty} \mathbf{E}[e^{-\theta_0 Q_{\tau \wedge n}}] = e^{-\theta_0 D}. \tag{4.45}$$

Finally, using Borel-Cantelli Lemma we can show τ is finite with probability one, which allows us to decompose (4.45) and characterize the boundary crossing probabilities as

$$\begin{aligned} \mathbf{Pr}(Q_\tau = 0) \cdot 1 + \mathbf{Pr}(Q_\tau = b) \cdot e^{-\theta_0 b} &= e^{-\theta_0 D}, \\ \mathbf{Pr}(Q_\tau = 0) + \mathbf{Pr}(Q_\tau = b) &= 1. \end{aligned}$$

Solving the above equations gives

$$\mathbf{Pr}(Q_\tau = 0) = \frac{e^{-\theta_0 D} - e^{-\theta_0 b}}{1 - e^{-\theta_0 b}}, \tag{4.46}$$

$$\mathbf{Pr}(Q_\tau = b) = \frac{1 - e^{-\theta_0 D}}{1 - e^{-\theta_0 b}}. \tag{4.47}$$

Taking the limit as $b \to \infty$ proves the claim. ■

Proof of Theorem 4.4.1. We first characterize the interruption probability for the cases $D > T$ and $D < T$ given $p(T)$, which is the interruption probability starting from $D = T$.

For any $x \geq 0$, define τ_x as the first hitting time of boundary x, i.e.,

$$\tau_x = \inf\{t \geq 0 : Q_t = x\}. \tag{4.48}$$

For the case $D > T$, using path-continuity of Q_t, strong Markov property and Lemma 4.4.1, we have

$$\begin{aligned} p(D) &= \mathbf{Pr}(\tau_0 < \infty | Q_0 = D) \\ &= \mathbf{Pr}(\tau_0 < \infty | Q_0 = T) \cdot \mathbf{Pr}(\tau_T < \infty | Q_0 = D) \\ &= e^{-\theta_0(D-T)} p(T). \end{aligned}$$

For the case $D < T$, we use the boundary crossing probabilities (4.46) and (4.47) that we derived in the proof of Lemma 4.4.1. Note that for the threshold policy when $D < T$, the drift is set to $R_1 - 1 = 2\theta_1$. Hence, by total probability theorem and strong Markov property, we obtain

$$\begin{aligned} p(D) &= \mathbf{Pr}(\tau_0 < \infty | Q_0 = D) \\ &= 1 \cdot \mathbf{Pr}(\tau_0 < \tau_T | Q_0 = D) + \mathbf{Pr}(\tau_0 < \infty | Q_0 = T) \cdot \mathbf{Pr}(\tau_T < \tau_0 | Q_0 = D) \\ &= \frac{e^{-\theta_1 D} - e^{-\theta_1 b}}{1 - e^{-\theta_1 b}} + p(T) \frac{1 - e^{-\theta_1 D}}{1 - e^{-\theta_1 b}}. \end{aligned}$$

We may obtain the desired result after simple manipulations of the above relation, once we compute $p(T)$.

In order to characterize $p(T)$, we use an analogue of one-step deviation analysis for Markov chains. Let $Q_0 = T$, and consider a small deviation Q_h, where h is a small time-step. Since Q_t is a Brownian motion with drift, Q_h has a normal distribution with variance h, and mean of $T + \alpha h$, where $\alpha \in [R_0 - 1, R_1 - 1]$. Therefore, the

probability of $Q_h \geq T$ is $(\frac{1}{2}+\delta)+o(h)$, where δ is a small constant of the same order of h, and $\frac{o(h)}{h} \to 0$ as $h \to 0$. By strong Markov property of the Brownian motion, (4.46) and (4.47), we have

$$\begin{aligned}
p(T) &= \mathbf{Pr}(\tau_0 < \infty | Q_0 = T) \\
&= \mathbf{Pr}(\tau_0 < \infty | Q_h \geq T)(\frac{1}{2}+\delta) + \mathbf{Pr}(\tau_0 < \infty | Q_h < T)(\frac{1}{2}-\delta) + o(h) \\
&= \Big[0 + p(T)\mathbf{E}_{Q_h}[\mathbf{Pr}(\tau_T < \infty)|Q_h \geq T]\Big](\frac{1}{2}+\delta) \\
&\quad + \Big[1 \cdot \mathbf{E}_{Q_h}[\mathbf{Pr}(\tau_0 < \tau_T)|Q_h < T] + p(T)\mathbf{E}_{Q_h}[\mathbf{Pr}(\tau_T < \tau_0)|Q_h < T]\Big](\frac{1}{2}-\delta) \\
&\quad + o(h) \\
&= \Big[0 + p(T)\mathbf{E}[e^{-\theta_0 Z}|Z \geq 0]\Big](\frac{1}{2}+\delta) \\
&\quad + \mathbf{E}\Big[\frac{e^{-\theta_1 T}(e^{-\theta_1 Z}-1)}{1-e^{-\theta_1 T}}\Big|Z < 0\Big](\frac{1}{2}-\delta) \\
&\quad + p(T)\mathbf{E}\Big[1 - \frac{e^{-\theta_1 T}(e^{-\theta_1 Z}-1)}{1-e^{-\theta_1 T}}\Big|Z < 0\Big](\frac{1}{2}-\delta) + o(h), \qquad (4.49)
\end{aligned}$$

where $Z = Q_h - T$ is a Gaussian random variable with mean αh and variance h. In order to obtain $p(T)$, we need to compute $E[e^{-\theta_0 Z}|Z \geq 0]$ and $E[e^{-\theta_1 Z}|Z < 0]$. We may compute these expressions exactly, but it is simpler to compute upper and lower bounds and then take the limit as $h \to 0$. By (4.42), we have

$$\begin{aligned}
E[e^{-\theta_0 Z}|Z \geq 0] &\leq \mathbf{E}[1 - \theta_0 Z + \frac{(\theta_0 Z)^2}{2}|Z \geq 0] = 1 - \theta_0\beta\sqrt{h} + o(\sqrt{h}), \\
E[e^{-\theta_0 Z}|Z \geq 0] &\geq \mathbf{E}[1 - \theta_0 Z|Z \geq 0] = 1 - \theta_0\beta\sqrt{h} + o(\sqrt{h}), \qquad (4.50)
\end{aligned}$$

where β is a constant. Similarly, we get

$$1 + \theta_1\beta\sqrt{h} + o(\sqrt{h}) \leq E[e^{-\theta_1 Z}|Z < 0] \leq 1 + \theta_1\beta\sqrt{h} + o(\sqrt{h}).$$

Plugging these relations back in (4.49), dividing by $\beta\sqrt{h}$ and taking the limit as h goes to zero, we obtain the following equation

$$p(T)\Big[\theta_0 + \frac{e^{-\theta_1 T}}{1-e^{-\theta_1 T}}\theta_1\Big] = \frac{e^{-\theta_1 T}}{1-e^{-\theta_1 T}}\theta_1, \qquad (4.51)$$

which gives the desired result for $p(T)$ after rearranging the terms. ■

Proof of Theorem 4.4.2. The proof technique for this theorem is analogue to that of Theorem 4.4.1. First, we consider the cases $D > T$ and $D < T$ and characterize the expected cost in terms of $J(T)$. Let τ_x be defined as in (4.48).

For the case $D > T$, note that no cost is incurred until the threshold T is reached. Hence

$$\begin{aligned}
J(D) &= \mathbf{E}\left[\int_0^{\tau_e} u_t dt \middle| Q_0 = D\right] \\
&= \mathbf{E}\left[\int_0^{\tau_e} u_t dt \middle| \tau_T = \infty, Q_0 = D\right] \mathbf{Pr}(\tau_T = \infty | Q_0 = D) \\
&\quad + \mathbf{E}\left[\int_0^{\tau_e} u_t dt \middle| \tau_T < \infty, Q_0 = D\right] \mathbf{Pr}(\tau_T < \infty | Q_0 = D) \\
&= 0 + \mathbf{E}\left[0 + \int_{\tau_T}^{\tau_e} u_t dt \middle| \tau_T < \infty, Q_0 = D\right] \mathbf{Pr}(\tau_T < \infty | Q_0 = D) \\
&\overset{(a)}{=} \mathbf{E}\left[\int_0^{\tau_e} u_t dt \middle| Q_0 = T\right] \mathbf{Pr}(\tau_T < \infty | Q_0 = D) \\
&= J(T)\mathbf{Pr}(\tau_T < \infty | Q_0 = D) \overset{(b)}{=} J(T) e^{-\theta_0 (D-T)},
\end{aligned}$$

where (a) follows from the memoryless property of Brownian motion and (b) is a consequence of Lemma 4.4.1.

For the case $D < T$, we can use a strong Markov property to write the following for a small time-step h:

$$\begin{aligned}
J(D) &= J(Q_0) = 1 \cdot h + \mathbf{E}_W[J(Q_h)] \\
&= h + \mathbf{E}_W\left[J(D) + \frac{\partial J}{\partial D}((R_1 - 1)h + W_h) + \frac{1}{2} \cdot \frac{\partial^2 J}{\partial D^2} \cdot h\right] + o(h) \\
&= h + J(D) + (R_1 - 1)\frac{\partial J}{\partial D} \cdot h + \frac{1}{2} \cdot \frac{\partial^2 J}{\partial D^2} \cdot h + o(h),
\end{aligned}$$

which gives the following ordinary differential equation after dividing by h and taking the limit as $h \to 0$

$$\frac{\partial^2 J}{\partial D^2} + \theta_1 \frac{\partial J}{\partial D} + 2 = 0, \quad 0 \le D \le T. \tag{4.52}$$

It is straightforward to solve the differential equation in (4.52) with the boundary condition $J(0) = 0$, and $J(T)$ as a parameter. This result completes the characterization of $J(D)$ described in (4.29) as a function of $J(T)$. Also, note that if we set the boundary condition $J(T) = 0$, $J(D)$ gives the expected time to hit either of the boundaries at 0 or T, i.e., we get

$$\mathbf{E}\Big[\min\{\tau_0, \tau_T\}\big|Q_0 = D\Big] = \frac{2}{\theta_1}\left[T \cdot \frac{1 - e^{-\theta_1 D}}{1 - e^{-\theta_1 T}} - D\right] \tag{4.53}$$

Now, we use a similar technique as in the proof of Theorem 4.4.1 to compute $J(T)$. Consider a small time step deviation $h > 0$ from the initial condition $Q_0 = T$. Similarly to (4.49), we have

$$\begin{aligned}
J(T) &= \gamma h + \Big[J(T)\mathbf{E}_{Q_h}[\mathbf{Pr}(\tau_T < \infty)|Q_h \geq T]\Big]\mathbf{Pr}(Q_h \geq T|Q_0 = T) \\
&\quad + \Big[1 \cdot \mathbf{E}\Big[\min\{\tau_0, \tau_T\}\big|Q_h < T\Big] + 0 \cdot \mathbf{E}_{Q_h}[\mathbf{Pr}(\tau_0 < \tau_T)|Q_h < T] \\
&\qquad + J(T)\mathbf{E}_{Q_h}[\mathbf{Pr}(\tau_T < \tau_0)|Q_h < T]\Big]\mathbf{Pr}(Q_h < T|Q_0 = T) + o(h) \\
&= \gamma h + J(T)\mathbf{E}[e^{-\theta_0 Z}|Z \geq 0](\frac{1}{2} + \delta) \\
&\quad + \mathbf{E}\Big[\min\{\tau_0, \tau_T\}\big|Q_h < T\Big](\frac{1}{2} - \delta) \\
&\quad + J(T)\mathbf{E}\Big[1 - \frac{e^{-\theta_1 T}(e^{-\theta_1 Z} - 1)}{1 - e^{-\theta_1 T}}\Big|Z < 0\Big](\frac{1}{2} - \delta) + o(h),
\end{aligned}$$

where γ is a constant bounded by 1, $\delta = \Theta(h)$, and $Z = Q_h - T$ is a Gaussian random variable with mean αh and variance h for some constant α. The second inequality in the preceding relations follows from (4.47) and Lemma 4.4.1. By (4.53), applying the bounds in (4.50) and (4.51), dividing by $\beta\sqrt{h}$ and taking the limit at $h \to 0$, we obtain the following equation

$$J(T)\left[\theta_1 \frac{e^{-\theta_1 T}}{1 - e^{-\theta_1 T}} + \theta_0\right] = \frac{2}{\theta_2}\left[1 - \theta_1 T \frac{e^{-\theta_1 T}}{1 - e^{-\theta_1 T}}\right],$$

which gives us the desired expression in (4.30) after some manipulations. ■

Proof of Theorem 4.4.3. In order to facilitate verification of the HJB equation (4.25), we rewrite and slightly manipulate the candidate solution $V(Q,p)$ given by (4.31). Recall that

$$V(Q,p) = \begin{cases} V_0(Q,p), & Q \geq T(Q,p) \\ V_1(Q,p), & Q \leq T(Q,p), \end{cases} \tag{4.54}$$

where

$$V_0(Q,p) = e^{-\theta_0(Q-T(Q,p))} J(T(Q,p)), \tag{4.55}$$

$$V_1(Q,p) = \left[J(T(Q,p)) + \frac{2}{\theta_1} T(Q,p)\right] \frac{1-e^{-\theta_1 Q}}{1-e^{-\theta_1 T(Q,p)}} - \frac{2}{\theta_1} Q, \tag{4.56}$$

$$J(T(Q,p)) = \frac{\frac{2}{\theta_1^2}\left[1-(1+\theta_1 T(Q,p))e^{-\theta_1 T(Q,p)}\right]}{\frac{\theta_0}{\theta_1} + (1-\frac{\theta_0}{\theta_1})e^{-\theta_1 T(Q,p)}}. \tag{4.57}$$

Note that $p^{T(Q,p)}(Q) = p$; and by definition of $p^T(\cdot)$ in (4.28) we may verify that the condition $Q \gtrless T(Q,p)$ is equivalent to $p \gtrless \frac{e^{-\theta_1 Q}}{\frac{\theta_0}{\theta_1}+(1-\frac{\theta_0}{\theta_1})e^{-\theta_1 Q}}$. Therefore, we can partition the feasible region $\mathcal{R}$ into two sub-regions $\mathcal{R}_0$ and $\mathcal{R}_1$, such that

$$\begin{aligned} \mathcal{R}_0 &= \left\{(Q,p) : p \geq \frac{e^{-\theta_1 Q}}{\frac{\theta_0}{\theta_1} + (1-\frac{\theta_0}{\theta_1})e^{-\theta_1 Q}}\right\} \cap \mathcal{R}, \\ \mathcal{R}_1 &= \left\{(Q,p) : p < \frac{e^{-\theta_1 Q}}{\frac{\theta_0}{\theta_1} + (1-\frac{\theta_0}{\theta_1})e^{-\theta_1 Q}}\right\} \cap \mathcal{R}. \end{aligned}$$

Hence, we need to verify HJB for two regions separately, using the proper expression in (4.54).

In order to verify the HJB equation for the candidate solution (4.31), we also need to characterize the optimal value of the minimization problem in (4.25). First, we characterize the optimal solution pair $(u^*, \hat{p}^*)$ for any feasible state $(Q,p) \in \mathcal{R}$. Observe that the optimization problem in (4.25) can be decomposed into two smaller

problems:

$$u^*(Q,p) = \text{argmin}_{u\in\{0,1\}}\left\{u + \frac{\partial V}{\partial Q}(R_u - 1)\right\}, \tag{4.58}$$

$$\hat{p}^*(Q,p) = \text{argmin}_{\hat{p}}\left\{\frac{1}{2}\frac{\partial^2 V}{\partial p^2}(\hat{p})^2 + \frac{\partial^2 V}{\partial Q \partial p}\hat{p}\right\}. \tag{4.59}$$

The minimization problem in (4.59) is quadratic and hence convex in $\hat{p}$. So we can use first order optimality condition to get

$$\hat{p}^*(Q,p) = -\frac{\partial^2 V}{\partial Q \partial p}(Q,p)/\frac{\partial^2 V}{\partial p^2}(Q,p). \tag{4.60}$$

For the problem in (4.58), $u^*(Q,p) = 0$ is and only if

$$0 + \frac{\partial V}{\partial Q}(R_0 - 1) \leq 1 + \frac{\partial V}{\partial Q}(R_1 - 1),$$

or equivalently

$$\frac{\partial V}{\partial Q}(Q,p) \geq -\frac{1}{R_1 - R_0}. \tag{4.61}$$

Using chain rule and implicit function theorem, we can analytically calculate $\frac{\partial V}{\partial Q}$ from (4.54) to conclude that the condition in (4.61) holds if and only if $(Q,p) \in \mathcal{R}_0$. In other words, $u^*(Q,p) = 0$ for all $(Q,p) \in \mathcal{R}_0$ and $u^*(Q,p) = 1$ for all $(Q,p) \in \mathcal{R}_1$.

In summary, the HJB equation in (4.25) boils down to the following equations:

$$0 = \frac{\theta_0}{2}\frac{\partial V_0}{\partial Q} + \frac{1}{2}\frac{\partial^2 V_0}{\partial Q^2} - \frac{1}{2}\left(\frac{\partial^2 V_0}{\partial Q \partial p}\right)^2 / \frac{\partial^2 V_0}{\partial p^2}, \quad \text{for all } (Q,p) \in \mathcal{R}_0, \tag{4.62}$$

$$0 = 1 + \frac{\theta_1}{2}\frac{\partial V_1}{\partial Q} + \frac{1}{2}\frac{\partial^2 V_1}{\partial Q^2} - \frac{1}{2}\left(\frac{\partial^2 V_1}{\partial Q \partial p}\right)^2 / \frac{\partial^2 V_1}{\partial p^2}, \quad \text{for all } (Q,p) \in \mathcal{R}_1, \tag{4.63}$$

where $V_0(Q,p)$ and $V_1(Q,p)$ are given by (4.55) and (4.56), respectively. The verification of (4.62) and (4.63) is straightforward but tedious. We omit the details for brevity. We may simply use symbolic analysis tools such as Mathematica for this part. ■

Proof of Theorem 4.4.4. From the proof of Theorem 4.4.3, we have characterized

the optimal policy $\pi^* : \mathbb{R} \times [0,1] \to \{0,1\} \times \mathbb{R}$ as

$$\pi^*(Q,p) = \Big(u^*(Q,p), \hat{p}^*(Q,p)\Big),$$

where $u^*(Q,p) = 0$ if and only if $(Q,p) \in \mathcal{R}_0$ and $\hat{p}^*(Q,p)$ given by (4.60) can be explicitly computed as follows:

$$\begin{aligned}
\hat{p}^*(Q,p) &= -\theta_0 p, \quad \text{for all } (Q,p) \in \mathcal{R}_0, \\
\hat{p}^*(Q,p) &= -\left[\frac{(1-p)e^{\theta_1 Q}}{(1-pe^{\theta_1 Q})^2} - \frac{1}{1-pe^{\theta_1 Q}}\right] \Big/ \left[\frac{(1-p)(e^{2\theta_1 Q} - e^{\theta_1 Q})}{\theta_1(1-pe^{\theta_1 Q})^2} - \frac{2(e^{\theta_1 Q}-1)}{\theta_1(1-pe^{\theta_1 Q})}\right] \\
&= -\frac{\theta_1}{(1+p)e^{\theta_1 Q} - 2}, \qquad \text{for all } (Q,p) \in \mathcal{R}_1. \quad (4.64)
\end{aligned}$$

Moreover, the dynamics of the state process under the optimal control policy is given by:

$$\begin{aligned}
dQ_t^* &= (R_{u^*(Q_t^*, p_t^*)} - 1)dt + dW_t, \\
dp_t^* &= \hat{p}^*(Q_t^*, p_t^*) dW_t.
\end{aligned}$$

For the proof of the first claim, observe that for a given manifold $\mathcal{M}(D,\epsilon)$, we have

$$T(Q,p) = T(D,\epsilon), \quad \text{for all } (Q,p) \in \mathcal{M}(D,\epsilon). \tag{4.65}$$

This claim holds by definition of $\mathcal{M}(D,\epsilon)$ in (4.32), and the fact that $T(Q,p)$ is the unique solution of $p^T(Q) = p$. Next, we show that if $(Q_t^*, p_t^*) \in \mathcal{M}(D,\epsilon)$ for any $t \geq 0$, then after executing the optimal policy π^*, the state process stays on the manifold $\mathcal{M}(D,\epsilon)$. First, consider the case where $Q_t^* \geq T(Q_t^*, p_t^*) = T(D,\epsilon)$. In this case, $u^*(Q_t^*, p_t^*) = 0$ and $\hat{p}^*(Q_t^*, p_t^*) = -\theta_0 p_t^*$. We would like to show that the solution of the SDE $dp_t^* = -\theta_0 p_t^* dW_t$ coincides with the invariant manifold given by $\tilde{p}_t = e^{-\theta_0(Q_t^* - T(D,\epsilon))} p(T(D,\epsilon))$. By employing Itō's Lemma we can check that $e^{-\theta_0(Q_t^* - T(D,\epsilon))} p(T(D,\epsilon))$ is indeed the desired solution. In particular, using the proper

evolution of the queue-length process Q_t^*, we can write

$$\begin{aligned} d\tilde{p}_t &= -\theta_0 e^{-\theta_0(Q_t^* - T(D,\epsilon))} p(T(D,\epsilon))(dQ_t^*) \\ &\quad + \frac{1}{2}\theta_0^2 e^{-\theta_0(Q_t^* - T(D,\epsilon))} p(T(D,\epsilon))dt \\ &= e^{-\theta_0(Q_t^* - T(D,\epsilon))} p(T(D,\epsilon))\Big[-\theta_0(\frac{\theta_0}{2}dt + dW_t) + \frac{1}{2}\theta_0^2\Big] \\ &= -\theta_0 e^{-\theta_0(Q_t^* - T(D,\epsilon))} p(T(D,\epsilon))dW_t \\ &= -p_t^* dW_t. \end{aligned}$$

Next, consider the case where $Q_t^* < T(Q_t^*, p_t^*) = T(D,\epsilon)$. In this case, $u^*(Q_t^*, p_t^*) = 1$ and $\hat{p}^*(Q_t^*, p_t^*)$ is given by (4.64). Similarly to the previous case, we may use Itō's Lemma to verify that the state process stays on the invariant manifold given by

$$\tilde{p}_t = p^{T(D,\epsilon)}(Q_t^*) = e^{-\theta_1 Q_t^*} + p(T(D,\epsilon))(1 - \frac{\theta_0}{\theta_1})\Big(1 - e^{-\theta_1 Q_t^*}\Big).$$

By Itō's Lemma we have

$$\begin{aligned} d\tilde{p}_t &= \frac{\partial p^{T(D,\epsilon)}(Q_t^*)}{\partial Q} dQ_t^* + \frac{1}{2}\cdot\frac{\partial^2 p^{T(D,\epsilon)}(Q_t^*)}{\partial Q^2}dt \\ &= -\theta_1 e^{-\theta_1 Q_t^*}\Big[1 - p(T(D,\epsilon))(1 - \frac{\theta_0}{\theta_1})\Big](\frac{\theta_1}{2}dt + dW_t) \\ &\quad + \frac{1}{2}\cdot\theta_1^2 e^{-\theta_1 Q_t^*}\Big[1 - p(T(D,\epsilon))(1 - \frac{\theta_0}{\theta_1})\Big]dt \\ &= -p_t^* dW_t = dp_t^*, \end{aligned}$$

which completes the proof of the first claim.

Now that we have established that the state process starting from (D,ϵ) under optimal control stays on a one-dimensional invariant manifold $\mathcal{M}(D,\epsilon)$, the optimality of the threshold policy $\pi^{T(D,\epsilon)}(Q)$ is immediate. Recall that the decision process of importance is $u_t^* \in \{0,1\}$, and we know that $u^*(Q,p) = 0$ if and only if $Q \geq T(Q,p)$. Moreover, since the optimal state process stays on $\mathcal{M}(D,\epsilon)$, we have $T(Q_t^*, p_t^*) = T(D,\epsilon)$. Hence, the optimal control policy (given the initial condition) chooses the action $u^*(Q,p) = 0$ if and only if $Q \geq T(D,\epsilon)$. Therefore, the

optimal policy $\pi^*(Q,p)$ coincides with the threshold policy $\pi^{T(D,\epsilon)}(Q)$. We may also verify that the interruption probability under the threshold policy conditioned on the history up to time t is given by p_t^*. ■

Chapter 5

Reliability Value of Energy Storage in Volatile Environments

In the last three chapters, we focused on systems motivated by media streaming applications in volatile environments such as wireless channels. We designed various control policies to optimize the quality, reliability or cost of media streaming for the end-user, while paying special attention to transient metrics.

In this chapter, we extend the analytical tools developed for media streaming applications to the context of power distribution systems in a volatile environment. Even though the system setup in power systems may appear vastly different than communication systems, there are many common elements. In the context of communication systems, packet delivery is negatively affected by channel unreliability and jitter, while in power systems supply and demand are subject to exogenous shocks. Untimely delivery of packets to a client results in playback interruption and degrades quality of experience. Similarly, late response to supply or demand shocks in power systems leads to a blackout event and degrades user experience. Packet buffering seems effective in reducing the effect of channel volatility in the context of media streaming applications. There is a simple but fundamental trade-off between the buffer size and playback interruptions. In an analogous manner, fast-response energy storages may be useful in reducing the negative effect of supply/demand volatility by masking some of the shocks to avoid blackouts. The goal of this part is to un-

derstand the effect of energy storage on the cost and distribution of blackouts in a volatile environments, and design proper control policies for management of the energy storage.

5.1 System Model

We examine an abstract model of system consisting of a single consumer, a single fully controllable supplier, a supplier with stochastic output (e.g., wind), and a storage system with finite capacity (Figure 5-1). These agents each represent an aggregate of several small consumers and producers. The details of the model are outlined below.

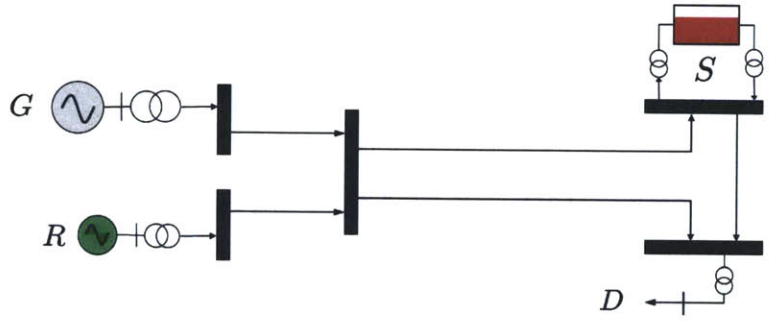

Figure 5-1: Layout of the physical layer of a power supply network with conventional and renewable generation, storage, and demand.

5.1.1 Supply

Controllable Supply

The controllable supply process is denoted by $\mathbf{G} = \{G_t : t \geq 0\}$, where G_t is the power output at time $t \geq 0$. It is assumed that the supplier's production is subject to an upward ramp constraint, in the sense that its output cannot increase instantaneously,

$$\frac{G_t - G_{t'}}{t - t'} \leq \zeta, \qquad \forall t : 0 \leq t < t'.$$

We do not assume a downward ramp constraint or a maximum capacity constraint on G_t. Thus, production can shut down instantaneously, and can meet any large

demand sufficiently far in the future.

Renewable Supply

The renewable supply process is denoted by $\mathbf{R} = \{R_t : t \geq 0\}$. It is assumed that $\mathbf{R}$ can be modeled as a process with two components: $\mathbf{R} = \overline{\mathbf{R}} + \Delta\mathbf{R}$, where $\overline{\mathbf{R}} = \{\overline{R}_t : t \geq 0\}$ is a deterministic process representing the predicted renewable supply, and $\Delta\mathbf{R} = \{\Delta R_t : t \geq 0\}$ is the residual supply assumed to be a random arrival process. Thus, at any given time $t \geq 0$, the total forecast supply from the renewable and controllable generators is given by $G_t + \overline{R}_t$.

5.1.2 Demand

The demand process is denoted by $\mathbf{D} = \{D_t : t \geq 0\}$, where D_t is the total power demand at time t, assumed to be exogenous and inelastic. Similar to the renewable supply, $\mathbf{D}$ has two components: $\mathbf{D} = \overline{\mathbf{D}} + \Delta\mathbf{D}$, where $\overline{\mathbf{D}} = \{\overline{D}_t : t \geq 0\}$ is the predicted demand process (deterministic), and $\Delta\mathbf{D} = \{\Delta D_t : t \geq 0\}$ is the residual demand, again, assumed to be a random arrival process.

Definition 5.1.1. The *power imbalance* is defined as the residual demand minus the residual supply.

$$P_t = \Delta D_t - \Delta R_t \tag{5.1}$$

The *normalized energy imbalance* is defined as:

$$W_t = \frac{P_t^2}{2\zeta} \tag{5.2}$$

5.1.3 Storage

The storage process is denoted by $\mathbf{s} = \{s_t \in [0, \overline{s}] : t \geq 0\}$, where s_t is the amount of stored energy at time t, and $\overline{s} < \infty$ is the storage capacity. The storage technology is subject to an upward ramp constraint:

$$\frac{s_t - s_{t'}}{t - t'} \leq r, \qquad \forall t : 0 \leq t < t'.$$

Thus, storage cannot be filled up instantaneously, though, it can be drained (to supply power) instantaneously. Let $\mathbf{U} = \{U_t : t \geq 0\}$, be the power withdrawal process from storage. The dynamics of storage is then given by:

$$s_t = s_0 + \int_0^t \mathbb{I}_{\{s_\tau < \bar{s}\}} r d\tau - \int_0^t U_\tau d\tau \tag{5.3}$$

It is desired to design a causal controller K such that the control law $U_t = K(s_t, G_t + R_t - D_t)$ maximizes the system reliability objectives.

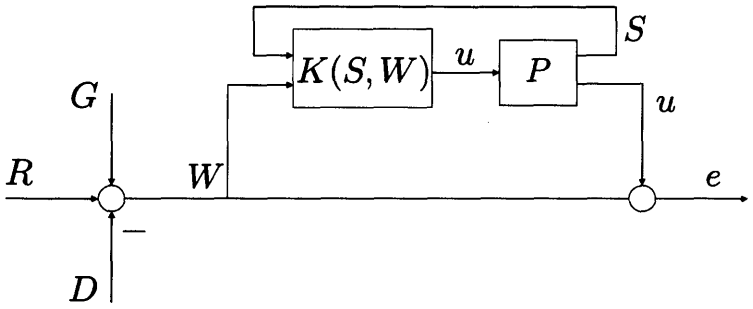

Figure 5-2: The control layer of the power supply network in Figure 5-1.

5.1.4 Reliability Metric

We refer to the event of not meeting the demand as a *blackout*. The cost of blackouts (COB) metric is defined as the expected long-term discounted cost of blackouts:

$$C^{\text{bo}} = \mathbf{E}\left[\int_0^\infty e^{-\theta\tau} h\left([P_\tau]^+\right) d\tau\right] \tag{5.4}$$

where $P(\cdot)$ is the power imbalance process, $h : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is an increasing function, and $\theta > 0$ is the discount rate.

5.1.5 Problem Formulation

In this section, we present the problem formulation. Before we proceed, we pose the following assumptions.

Assumption 5.1.1. *The normalized energy imbalance process* (5.2) *is the jump process in a compound poisson process with arrival rate Q and jump size distribution f_W,*

where the support of f_W lies within a bounded interval $[0, B]$. The maximum jump size is thus upperbounded by B.

Assumption 5.1.2. *The forecast supply is equal to the forecast demand. That is:*

$$\overline{D}_t = G_t + \overline{R}_t, \qquad \forall t \geq 0$$

Under Assumption 5.1.2, the energy from storage will be used only to compensate for the power imbalance, since in the absence of an *energy* shock, supply is equal to demand, and storage provides no additional utility. Under Assumptions 5.1.1 and 5.1.2, the dynamics of the storage process can be written as:

$$s_t = s_0 + \int_0^t \mathbb{I}_{\{s_\tau < \overline{s}\}} r d\tau - \int_0^t \mu\left(s_{\tau^-}, W_\tau\right) dN_\tau \tag{5.5}$$

where N_t is a Poisson process of rate Q, and W_t is the jump size (energy imbalance) process, drawn independently and identically from a distribution f_W. Further, μ denotes a *control policy*. We focus on stationary Markov policies since the energy imbalance modeled as a compound Poisson process is stationary and memoryless. We denote the set of all such feasible policies by Π.

We are now ready to state the problem formulation. Let $C_\mu(s)$ denote the expected long-term discounted cost of blackouts starting from an initial state s and under control policy μ,

$$C_\mu(s) = \mathbf{E}\left[\sum_{k=1}^{\infty} e^{-\theta t_k} g\left(W_k - \mu(s_{t_k^-}, W_k)\right) \middle| s_0 = s\right], \tag{5.6}$$

where t_k is the k-th Poission arrival time, and $W_k = W_{t_k}$ is size of the k-th jump. Moreover, $g : [0, B] \to \mathbb{R}$ is the stage cost as a function of energy imbalance (blackout size). In this work, we assume the following assumptions hold.

Assumption 5.1.3. *The stage cost function $g(\cdot)$ is bounded, strictly increasing and continuously differentiable. Moreover, $\mathbf{E}_W[g(W)] > 0$, and $g(0) = 0$.*

The system reliability problem can now be formulated as an infinite horizon

stochastic optimal control problem

$$C_\mu(s) \to \min_{\mu \in \Pi} \tag{5.7}$$

where the optimization problem (5.7) is subject to the state dynamics (5.5). A policy $\mu^* \in \Pi$ is defined to be optimal if

$$\mu^* \in \arg\min_{\mu \in \Pi} C_\mu(s).$$

The associated *value function* or optimal cost function is denoted by $C(s)$, where

$$C(s) = \min_{\mu \in \Pi} C_\mu(s), \quad 0 \leq s \leq \bar{s}. \tag{5.8}$$

5.2 Main Results

5.2.1 Characterizations of the Value Function

We first provide several characterizations for the value function defined in (5.8) and establish specific properties that are useful in characterization of the optimal policy.

Let $J_\mu(s, w)$ be the expected long-term discounted cost under policy μ conditioned on the first jump arriving at time $t_1 = 0$, and being of size w. Here, s is the state of the system before executing the action dictated by the policy. By the memoryless property of the Poisson process, we have

$$\begin{aligned} J_\mu(s, w) &= g(w - \mu(s, w)) \\ &+ \mathbf{E}\Big[\sum_{k=1}^{\infty} e^{-\theta t_k} g(W_k - \mu(s_{t_k^-}, W_k)) \Big| s_0 = s - \mu(s, w)\Big] \end{aligned} \tag{5.9}$$

We may relate $J_\mu(s, W)$ to the total expected cost $C_\mu(s)$ defined in (5.6) as follows:

$$C_\mu(s) = \mathbf{E}\left[e^{-\theta t_0} J_\mu(\min\{s + r t_0, \bar{s}\}, W)\right], \tag{5.10}$$

where t_0 is an exponential random variable with mean $1/Q$, and is independent of

W, drawn from distribution f_W.

From (5.10), it is clear that from the minimization of J_μ across all admissible policies Π, we may obtain the optimal solution to the original problem in (5.8). The discrete-time formulation of J_μ given by (5.9), facilitates deriving the Bellman equation as the necessary and sufficient optimality condition, as well as development of efficient numerical methods. We summarize these results in the following theorem.

Theorem 5.2.1. *Given an admissible control policy $\mu \in \Pi$, let $J_\mu : [0, \bar{s}] \times [0, B] \mapsto \mathbb{R}$ be the function defined as in (5.9). A function $J : [0, \bar{s}] \times [0, B] \mapsto \mathbb{R}$ satisfies*

$$J(s, w) = J^*(s, w) \stackrel{\text{def}}{=} \min_{\mu \in \Pi} J_\mu(s, w), \qquad \forall (s, w),$$

if and only if it satisfies the following fixed-point equation:

$$J(s, w) = (TJ)(s, w) \stackrel{\text{def}}{=} \min_{u \in [0, \min\{s, w\}]} \Big\{ g(w - u) + \mathbf{E}\Big[e^{-\theta t_0} J\big(\min\{s - u + r t_0, \bar{s}\}, W \big) \Big] \Big\}, \tag{5.11}$$

Moreover, a stationary policy $\mu^*(s, w)$ is optimal if and only if $u = \mu^*(s, w)$ achieves the minimum in (5.11) for $J = J^*$. Finally, the value iteration algorithm

$$J_{k+1} = T J_k, \tag{5.12}$$

converges to J^* for any initial condition J_0.

Proof. The result follows from establishing the contraction property of T, which is standard for discounted problems with bounded stage cost. See [52] for more details. □

An alternative approach to characterization of the optimal cost function is based on continuous-time analysis of problem (5.8), which leads to Hamilton-Jacobi-Bellman (HJB) equation. In the following theorem we present some basic properties of the optimal cost function as well as the HJB equation.

Theorem 5.2.2. *Let $C(s)$ be the optimal cost function defined in (5.8). The following statements hold:*

(i) *$C(s)$ is strictly decreasing in s.*

(ii) *If the stage cost $g(\cdot)$ is convex, the optimal cost function $C(s)$ is also convex in s.*

(iii) *If C is continuously differentiable, then for all $s \in [0, \bar{s}]$, it satisfies the following HJB equation*

$$\frac{dC}{ds} = \frac{Q+\theta}{r}C(s) - \frac{Q}{r}\mathbf{E}\Big[\min_{u\in[0,\min\{s,W\}]} g(W-u) + C(s-u)\Big], \tag{5.13}$$

with the boundary condition

$$\frac{dC}{ds}\Big|_{s=\bar{s}} = 0. \tag{5.14}$$

Moreover, the optimal policy $\mu^*(s,w)$ achieves the optimal solution of the minimization problem in (5.13). Furthermore, for a given policy μ, if the cost function $C_\mu(s)$ is differentiable, it satisfies the following delay differential equation

$$\frac{dC_\mu}{ds} = \frac{Q+\theta}{r}C_\mu(s) - \frac{Q}{r}\mathbf{E}\Big[g(W-\mu(s,W)) + C_\mu(s-\mu(s,W))\Big], \tag{5.15}$$

with the boundary condition given by (5.14).

Proof. See the Appendix. □

The result of Theorem 5.2.2 part (iii) requires continuous differentiability of the optimal cost function, which can be established under some mild conditions such as differentiability of the stage cost function g and the probability density function $f_W(\cdot)$ of Poisson jumps (cf. Benveniste and Scheinkman [53]). Throughout this chapter, we assume that $C(s)$ is in fact continuously differentiable and the results of Theorem 5.2.2 are applicable.

5.2.2 Characterizations of the Optimal Policy

In this part, we derive some structural properties of the optimal policy using the optimal cost characterizations given in Theorems 5.2.1 and 5.2.2. First, we show that the myopic policy of allocating reserve energy from storage to cover as much of every shock as possible is optimal for linear stage cost functions. Then, we partially characterize the structure of optimal policy for strictly convex stage cost functions.

Theorem 5.2.3. *If the stage cost is linear, i.e., $g(x) = \beta x$ for some $\beta > 0$, then the myopic policy*

$$\mu^*(s, w) = \min\{s, w\}, \tag{5.16}$$

is optimal for problem (5.8).

Proof. See the Appendix. □

Next, we focus on nonlinear but convex stage cost functions. In this case, the myopic policy defined in (5.16) is no longer optimal. Intuitively, the myopic policy greedily consumes the reserve and thereby increases the chance of a large blackout. In the linear stage cost case, the penalty for a large blackout is equivalent to the total penalty of many small blackouts. This is contrary to the strictly convex case. Therefore, the optimal policy in this case tends to be more conservative in consuming the reserve. Nevertheless, the structure of the optimal policy shows some similarities with the myopic policy. In the following we present some characterizations of the structural properties of the optimal policy using the results from Section 5.2.1.

Assumption 5.2.1. *The storage process has a positive drift in the sense that the rate of the compound Poisson process is less than the ramp constraint, i.e.,*

$$Q\mathbf{E}[W] \leq r.$$

Theorem 5.2.4. *Let $\mu^*(s, w)$ be the optimal policy associated with problem (5.8). If Assumption 5.2.1 holds, then $\mu^*(s, w)$ is monotonically nondecreasing in both s and w.*

Proof. See the Appendix. □

Theorem 5.2.5. *Let μ^* denote the optimal policy associated with problem (5.8) with strictly convex stage cost $g(\cdot)$. There exist a unique kernel function $\phi : [-B, \bar{s}] \to \mathbb{R}$ such that*

$$\mu^*(s,w) = \Big[w - \phi(s-w)\Big]^+, \quad \forall (s,w) \in [0,\bar{s}] \times [0,B], \tag{5.17}$$

where,

$$\begin{aligned} \phi(p) \;=\; & \arg\min_{x} \quad g(x) + C(x+p) \\ & \text{s.t.} \quad x \le \min\{B, \bar{s}-p\} \\ & \qquad x \ge \max\{0, -p\} \end{aligned} \tag{5.18}$$

Moreover, under Assumption 5.2.1, we can represent the kernel function $\phi(p)$ as follows:

$$\phi(p) = \begin{cases} -p, & -B \le p \le b_0 \\ \phi^\circ(p), & b_0 \le p \le b_1 \\ 0, & b_1 \le p \le \bar{s}, \end{cases} \tag{5.19}$$

where $\phi^\circ(p)$ is the unique solution of

$$g'(x) + C'(x+p) = 0, \tag{5.20}$$

and b_0 and b_1 are the break-points, where

$$b_0 = -(g')^{(-1)}\Big(C'(0)\Big) \ge -(g')^{(-1)}\Big(\frac{Q}{r}\mathbf{E}[g(W)]\Big) \ge -B, \tag{5.21}$$

$$b_1 = -(C')^{(-1)}\Big(g'(0)\Big) \le \bar{s}. \tag{5.22}$$

Proof. See Appendix. □

Theorem 5.2.5 demonstrates a very special structure for the optimal policy. In fact, it shows that the two dimensional policy can be represented using a single dimensional kernel function. This result allows us to significantly reduce the computational

complexity of numerical methods for computing the optimal policy. In addition, using Theorem 5.2.5, we can provide a qualitative picture of the structure of the optimal policy. Figures 5-3 and 5-4 illustrate a conceptual plot of the kernel function, and the optimal policy, respectively.

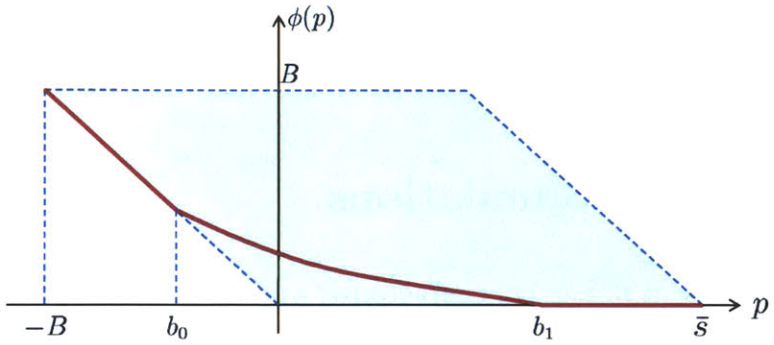

Figure 5-3: Structure of the kernel function $\phi(p)$ defined in (5.18).

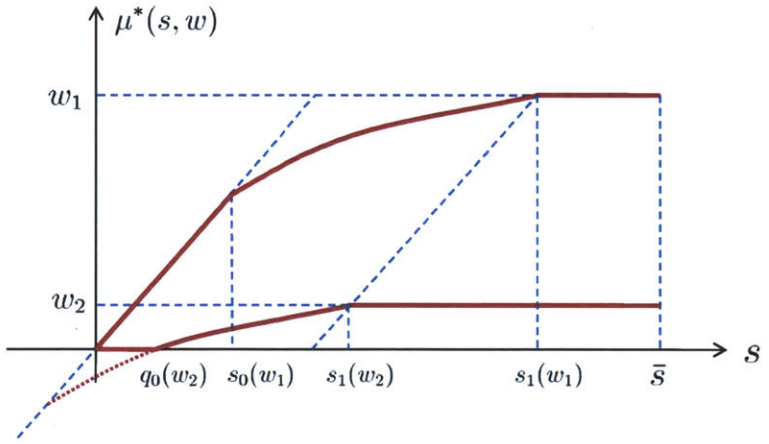

Figure 5-4: Structure of the optimal policy $\mu^*(s,w)$ for a convex stage cost, for $w = w_1, w_2$.

In particular, we can summarize the characterization of the optimal policy as follows. If $w \geq -b_0$, we have

$$\mu^*(s,w) = \begin{cases} s, & 0 \leq s \leq s_0(w) \\ w - \phi^\circ(s-w), & s_0(w) \leq s \leq s_1(w) \\ w, & s_1(w) \leq s \leq \bar{s}, \end{cases} \tag{5.23}$$

where $s_i(w) = w + b_i$ for $i = 0, 1$. In the case where $w \leq -b_0$, we have

$$\mu^*(s,w) = \begin{cases} 0, & 0 \leq s \leq q_0(w) \\ w - \phi^\circ(s-w), & q_0(w) \leq s \leq s_1(w) \\ w, & s_1(w) \leq s \leq \bar{s}, \end{cases} \tag{5.24}$$

where $q_0(w)$ is the unique solution of $\phi^\circ(s-w) = w$.

5.3 Numerical Simulations

In this part, we present numerical characterizations of the optimal cost function and optimal policy in different scenarios. Moreover, we study the effect of storage size and volatility on system performance, for various control policies.

We use the value iteration algorithm (5.12) to compute the optimal policy and cost function for nonlinear stage costs. Figures 5-5 and 5-6 illustrate the optimal policy and cost function in a scenario with uniformly distributed random jumps, quadratic stage cost, and the following parameters: $\theta = 0.1, r = 1, Q = 0.8, \bar{s} = 2$. Observe that the optimal policy complies with the conceptual Figure 5-4.

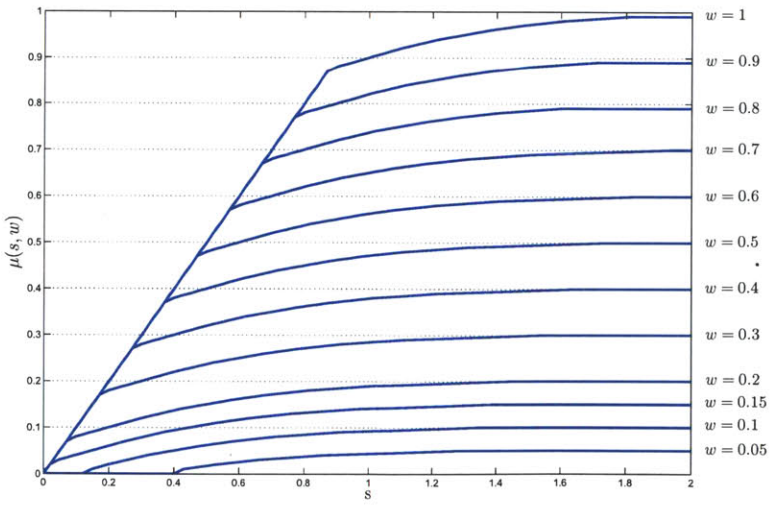

Figure 5-5: Optimal policy computed by value iteration algorithm (5.12) for quadratic stage cost and uniform shock distribution.

Figure 5-7 shows the value of storage, defined as the normalized improvement

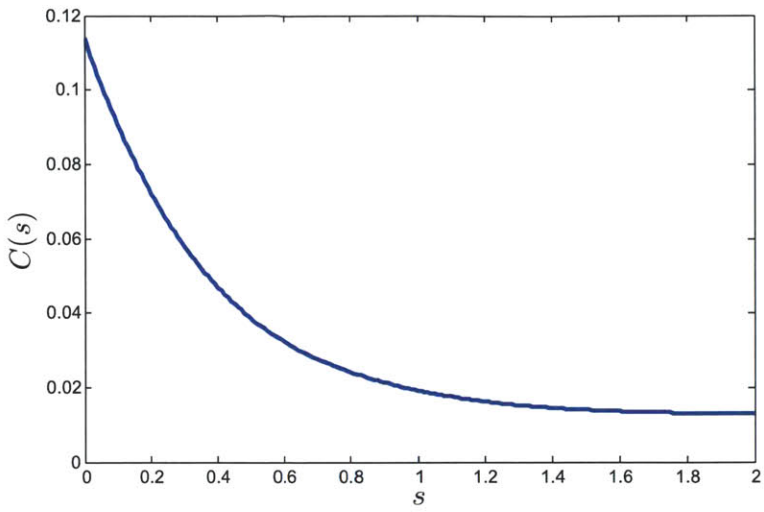

Figure 5-6: Optimal cost function computed by value iteration algorithm (5.12) for quadratic stage cost and uniform shock distribution.

of energy storage in expected cost, for different Poisson arrival rates. In this case $\theta = 0.01, g(x) = x^3, r = 1, W = 1$. Note that the storage process has a negative drift if and only if $Q > 1$. Observe that in the positive or zero drift cases, even a small value of storage yields a significant effect in reducing the blackout cost. However, in the negative drift case, the value of storage is significantly lower. Observe that for the negative drift case, there is a critical storage size that yields a sharp improvement in the value of storage.

5.3.1 Blackout Statistics

We discussed in Section 5.2.2 that the myopic policy given by (5.16) is not necessary optimal for nonlinear stage cost functions. In this part, we study the effect of different optimal policies, in the sense of (5.7), for different stage costs on the distribution of large blackouts. Figure 5-8 shows the blackout distribution in a scenario with deterministic jumps of size one, for both myopic policy and the optimal policy for a cubic cost function. Note that, the stage cost for the non-myopic policy assigns a significantly higher weight to larger blackouts. Therefore, as we can see in Figure 5-8, the non-myopic policy results in less frequent large blackouts at the price of more frequent small blackouts.

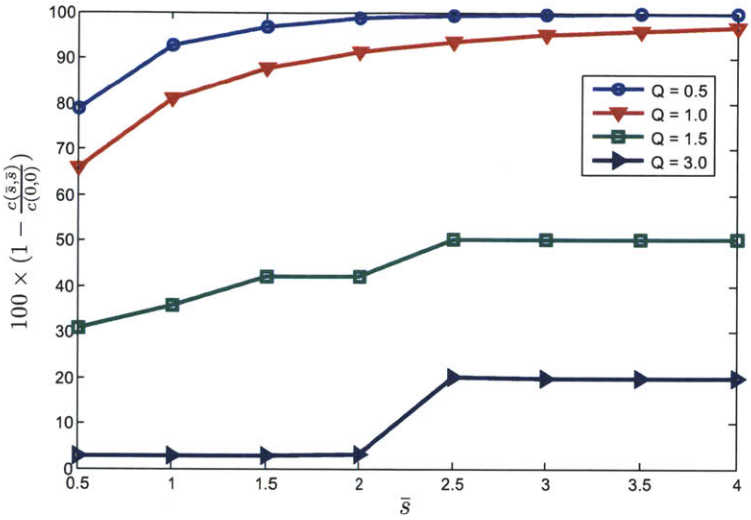

Figure 5-7: Value of energy storage as a function of the storage capacity for different Poisson arrival rates. $c(s;\bar{s})$ denotes the optimal cost function (5.8) when the storage capacity is given by $\bar{s}$.

Next, we study the effect of storage size on probability of large blackouts. Figure 5-9 plots this metric for different policies that are all optimal for different stage cost functions. Similarly to Figure 5-7, we observe a sharp improvement of the reliability metric at a critical storage size. It is worth mentioning that given a target reliability metric, the storage size required by the optimal policy with cubic stage cost is about half of what is required by the myopic policy.

Finally, we compare the reliability of myopic and non-myopic policies in terms of probability of large blackouts as a function of the volatility of the demand/supply process. We define volatility as the energy of the shock process, i.e.,

$$\text{volatility} = Q\mathbf{E}[W^2],$$

which depends both on the mean and variance of the compound arrival process. Figure 5-10 demonstrates large blackout probabilities as a function of volatility, for a system with uniformly distributed jumps with constant mean $R\mathbf{E}[W] = 1$. As shown in Figure 5-10, higher volatility increases the probability of large blackouts in an almost linear fashion.

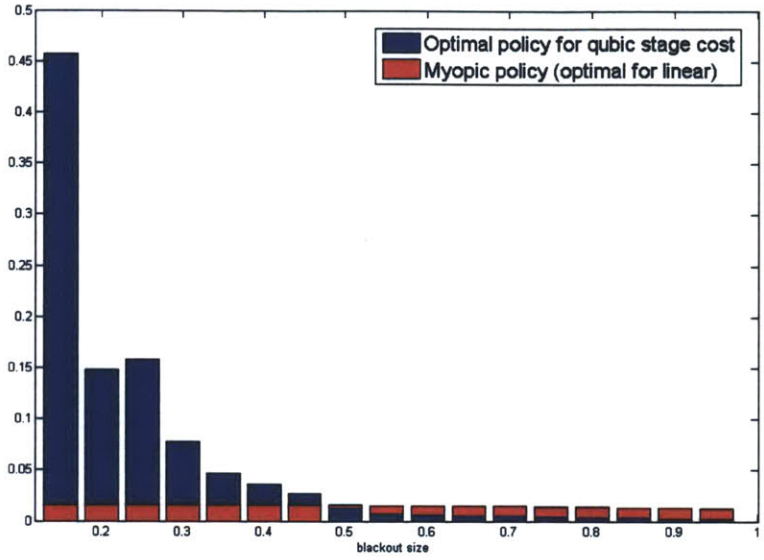

Figure 5-8: Blackout distribution comparison of myopic and non-myopic policies (deterministic jumps with rate $Q = 0.8$).

5.4 Appendix to Chapter 5 - Proofs

Proof of Theorem 5.2.2: *Part (i):* The monotonicity property of the value function follows almost immediately from the definition. Let $0 \leq s_1 < s_2 \leq \bar{s}$, and assume $C(s) = C_\mu(s)$ for some policy μ. Given the initial state s_1, let $u_t^{(1)}$ be the control process under policy μ. Note that for every realization ω of the compound Poisson process, the sample path $u_t^{(1)}(\omega)$ is admissible for initial condition $s_2 > s_1$. Therefore, by definitions (5.6) and (5.8), we have $C(s_2) \leq C(s_1)$.

In order to show the strict monotonicity, consider the controlled process starting from s_1. Let τ be the first arrival time such that $g(W_\tau - u_\tau^{(1)}) > 0$. By Assumption 5.1.3, we have $\mathbf{Pr}(\tau \in [0, T]) > 0$ for some $T < \infty$. For every sample path ω, define the control process

$$u_t^{(2)}(\omega) = u_t^{(1)}(\omega) + \delta \cdot \mathbb{I}_{\{t=\tau(\omega)\}},$$

for some $\delta > 0$ such that $\delta \leq \min\{s_2 - s_1, W_{\tau(\omega)} - u_{\tau(\omega)}^{(1)}\}$.

It is clear that $u_t^{(2)}(\omega)$ is admissible for the controlled process starting from s_2.

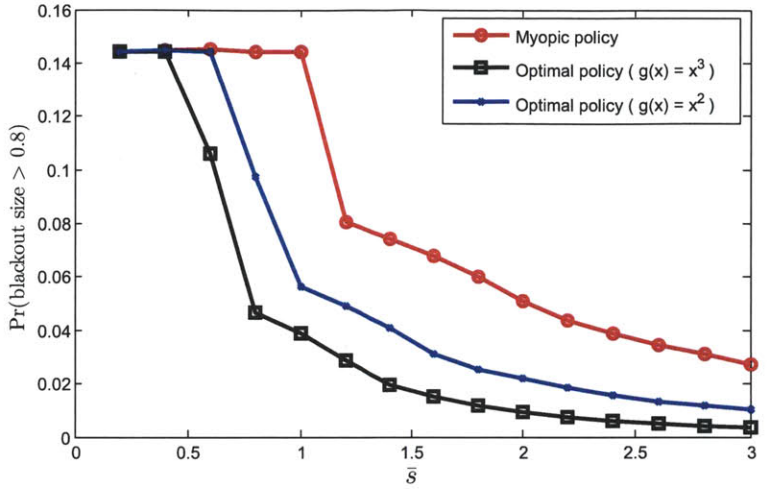

Figure 5-9: Probability of large blackouts as a function of storage size for different policies (deterministic jumps with rate $Q = 1.0$).

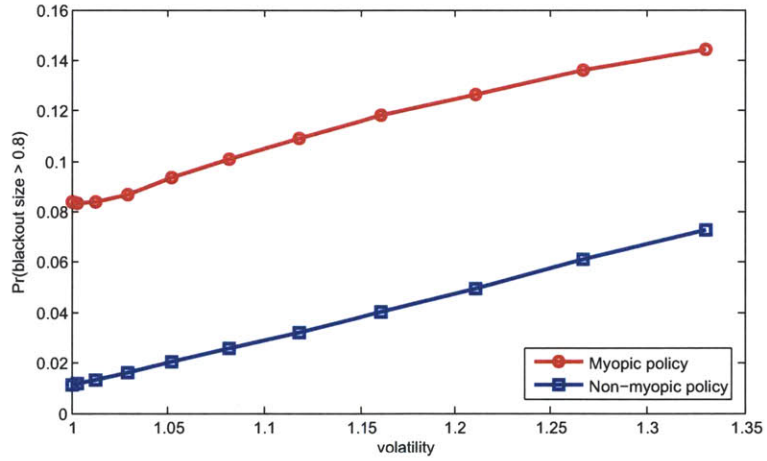

Figure 5-10: Probability of large blackouts vs. volatility for different policies (uniformly distributed random jumps with $Q = 1.0$ and $E[W] = 1$).

Using the definition of the expected cost function in (5.6), we can write

$$\begin{aligned} C(s_1) - C(s_2) &= \mathbf{E}_\omega[e^{-\theta\tau(\omega)} g(W_{\tau(\omega)} - u^{(1)}_{\tau(\omega)}) \\ &\quad -e^{-\theta\tau(\omega)} g(W_{\tau(\omega)} - u^{(1)}_{\tau(\omega)} - \delta)] \\ &\geq \mathbf{E}[\epsilon e^{-\theta\tau(\omega)}], \quad \text{for some } \epsilon > 0 \\ &\geq \epsilon e^{-\theta T} \mathbf{Pr}(\tau \in [0, T]) > 0, \end{aligned}$$

where the first inequality holds by strict monotonicity of g.

Part (ii): We first prove convexity of $J^*(s, w)$ defined in Theorem 5.2.1, and use

it to establish convexity of $C(s)$.

In order to show convexity of $J^*(s,w)$, we need to show that the operator T defined in (5.11) preserves convexity. Then the claim would be immediate using the convergence of value iteration algorithm (5.12) to optimal cost J^*, where the initial condition is an arbitrary convex function such as $J_0 = 0$.

Next we show that the operator T preserves convexity for this particular problem. Define the objective function in (5.11) as $Q(s,w,u)$. We have

$$\begin{aligned} Q(s,w,u) &= g(w-u) + \mathbf{E}\Big[e^{-\theta t_0} J\big(\min\{s-u+rt_0, \bar{s}\}, W\big)\Big] \\ &= g(w-u) + \int_{\frac{\bar{s}-s+u}{r}}^{\infty} e^{-\theta t_0}\mathbf{E}[J\big(\bar{s}, W\big)]Re^{-Qt_0}dt_0 \\ &+ \int_0^{\frac{\bar{s}-s+u}{r}} e^{-\theta t_0}\mathbf{E}[J\big(s-u+rt_0, W\big)]Re^{-Qt_0}dt_0. \end{aligned}$$

Using the fact that J is convex, linearity of expectation and basic definition of a convex function, it is straightforward but tedious to show that $Q(s,w,u)$ is a convex function. We omit the details for brevity. Given the convexity of Q, the convexity of $(TJ)(s,w)$ is immediate, since we are minimizing a multidimensional convex function over one of its dimensions. Hence, we have established convexity of $J^*(s,w)$ in (s,w). Finally, we can express $C(s)$ in terms of $J^*(s,w)$ as in (5.10). This results in convexity of $C(s)$ using the above argument for proving convexity of $Q(s,w,u)$.

Part (iii): The derivation of Hamilton-Jacobi-Bellman is relatively standard. We present a proof sketch based on *principle of optimality.* For a more detailed treatment, please refer to [52], [54] and [55].

Let s_t be the state process under the optimal policy governed by the SDE in (), and $s_0 = s < \bar{s}$. By principle of optimality, going from time 0 to time h, we have

$$C(s) \;=\; C(s_0) = \min_{u_t \in \Pi} \left\{ \mathbf{E}\Big[\int_0^h e^{-\theta\tau} g(W_\tau - u_\tau) dN_\tau\Big] + e^{-\theta h}\mathbf{E}\big[C(s_h)\big] \right\},$$

where the expectation is with respect to the compound Poisson process. Note that in this particular model, we assume that the control process u_t is progressively mea-

surable with respect to the jump process.

Let N_h be the number of Poisson arrivals in $[0, h]$, where h is positive but small. There are three cases to consider: First, there are no arrivals in $[0, h]$, in which case the decision function is trivially zero, and no penalty occurs. Second, there is a single arrival in this interval. In this case, the control function u_t is reduced to a scalar decision u that is measurable w.r.t jump size W. Third, there are more than one arrivals in $[0, h]$, which occurs with probability $o(h)$. Since, the stage cost is bounded, the expected cost under this condition remains $o(h)$. Hence, for every $s < \bar{s}$,

$$\begin{aligned} C(s) &= \mathbf{E}_W\Big[\min_{0\le u\le s, W}\Big\{0\cdot\mathbf{Pr}(N_h=0) \\ &\quad + g(W-u)\mathbf{Pr}(N_h=1) \\ &\quad + e^{-\theta h}C(s+rdh)\mathbf{Pr}(N_h=0) \\ &\quad + e^{-\theta h}C(s-u+rdh)\mathbf{Pr}(N_h=1)+o(h)\Big\}\Big] \\ &= e^{-\theta h}C(s)+\mathbf{E}_W\Big[\min_{0\le u\le s, W}\Big\{Rh\ g(W-u) \\ &\quad + e^{-\theta h}(1-Rh)(C(s+rdh)-C(s)) \\ &\quad + e^{-\theta h}Rh\ C(s-u)+o(h)\Big\}\Big]. \end{aligned} \tag{5.25}$$

Using the fact that C is differentiable on $[0, \bar{s})$, we may verify the result in (5.13), by dividing the above relation by h and taking the limit as h tends to zero.

The derivation of the boundary condition in (5.14) is similar. Note that for $s = \bar{s}$, when no Poisson arrival occurs in interval $[0, h]$, the state of the system stays at $\bar{s}$. Therefore, we can modify (5.25) as follows:

$$\begin{aligned} C(\bar{s}) = \mathbf{E}_W\Big[\min_{0\le u\le \bar{s}, W}\Big\{&0\cdot\mathbf{Pr}(N_h=0) \\ &+ g(W-u)\mathbf{Pr}(N_h=1)+e^{-\theta h}C(\bar{s})\mathbf{Pr}(N_h=0) \\ &+ e^{-\theta h}C(\bar{s}-u)\mathbf{Pr}(N_h=1)+o(h)\Big\}\Big]. \end{aligned} \tag{5.26}$$

Again, dividing by h and taking the limit as h goes to zero, we have

$$\frac{Q+\theta}{r}C(\bar{s}) - \frac{Q}{r}\mathbf{E}_W\Big[\min_{0\le u\le \bar{s}, W} g(W-u) + C(\bar{s}-u)\Big] = 0,$$

which is the same as (5.14) in light of (5.13).

The derivation of the differential equation (5.15) for a particular policy is similar, noting the memoryless property of Poisson process and stationarity of the controlled state process. ■

Proof of Theorem 5.2.3: We establish optimality of μ^* by showing that it achieves an expected cost no higher than any other admissible policy. Consider an admissible policy $\tilde{\mu}$ such that $\tilde{\mu}(s,w) < \min\{s,w\}$ for some $(s,w) \in [0,\bar{s}] \times [0,B]$. For every sample path of the controlled process, let $\tau_1(\omega)$ be the first Poisson arrival time such that

$$\min\{s_{\tau_1^-}, W_{\tau_1}\} - \tilde{\mu}(s_{\tau_1^-}, W_{\tau_1}) = \epsilon > 0.$$

Therefore, by applying policy $\tilde{\mu}$ instead of μ^*, we pay an extra penalty of $\beta\epsilon e^{-\theta\tau_1(\omega)}$. The reward for this extra penalty is that the state process is now biased by at most ϵ, which allows us to avoid later penalties. However, since the stage cost is linear, the penalty reduction by this bias for any time $\tau_2(\omega) > \tau_1(\omega)$ is at most $\beta\epsilon e^{-\theta\tau_2(\omega)}$. Hence, for this sample path ω, the policy $\tilde{\mu}$ does worse than the myopic policy μ^* at least by $\beta\epsilon(e^{-\theta\tau_1(\omega)} - e^{-\theta\tau_2(\omega)}) > 0$. Therefore, by taking the expectation for all sample paths, the myopic policy cannot do worse than any other admissible policy. Note that this argument does not prove the uniqueness of μ^* as the optimal policy. In fact, we may construct optimal policies that are different from μ^* on a set $A \subseteq [0,\bar{s}] \times [0,B]$, where $\mathbf{Pr}((s_{t^-}, W_t) \in A) = 0$. ■

We delay the proof of Theorem 5.2.4 until after proof of Theorem 5.2.5. Let us start with some useful lemmas on the structure of the kernel function.

Lemma 5.4.1. *Let $\phi(p)$ be defined as in (5.18). We have*

1. *If* $\phi(p_0) = -p_0$ *for some* p_0, *then*

$$\phi(p) = -p, \quad \textit{for all } p \leq p_0.$$

2. *If* $\phi(p_1) = 0$ *for some* p_1, *then*

$$\phi(p) = 0, \quad \textit{for all } p \geq p_1.$$

Proof. By convexity of the stage cost function and Theorem 5.2.2(ii), $\phi(p)$ is the optimal solution of a convex program. Therefore, if $\phi(p_0) = -p_0$ for some $p_0 \leq 0$, we have

$$g'(-p_0) + C'(0) \geq 0.$$

Thus, by convexity of stage cost, $g(-p) \geq g(-p_0)$, for any $p \leq p_0$. Therefore, by convexity of $C(\cdot)$ and $g(\cdot)$,

$$g'(x) + C'(x+p) \geq g'(-p) + C'(0) \geq 0, \text{for all } x \geq -p,$$

which immediately implies optimality of $(-p)$, for $p \leq p_0$.

Similarly, for the case where $\phi(p_1) = 0$, we have $g'(0) + C'(p_1) \geq 0$, which implies

$$g'(x) + C'(x+p) \geq g'(0) + C'(p) \geq 0, \quad \text{for all } p \geq p_1,$$

hence, the objective is nondecreasing for all feasible x and $\phi(p) = 0$. □

Lemma 5.4.2. *Let* $C(s)$ *be defined as in (5.8), and assume that the stage cost* $g(\cdot)$ *is convex. Then*

$$\frac{dC}{ds}(s) \geq -\frac{Q}{r}\mathbf{E}_W[g(W)], \quad 0 \leq s \leq \bar{s}. \tag{5.27}$$

Proof. By Theorem 5.2.2(ii), the optimal cost function $C(s)$ is convex. Hence, $\frac{dC}{ds}(s) \geq \frac{dC}{ds}(0)$. On the other hand, by Theorem 5.2.2(iii), we can write

$$\frac{dC}{ds}(0) \;=\; \frac{Q+\theta}{r}C(0) - \frac{Q}{r}\mathbf{E}_W\Big[\min_{u=0} g(W-0) + C(0)\Big].$$

Combining the two preceding relations proves the claim. □

Lemma 5.4.3. *If Assumption 5.2.1 holds, then the first constraint in (5.18) is never active, i.e.,* $\phi(p) < \min\{B, \bar{s} - p\}$.

Proof. We show that under Assumption 5.2.1, the slope of the objective function is always non-negative at $x = \min\{B, \bar{s} - p\}$. In the case where $\bar{s} - p \leq B$, we have

$$\frac{\partial}{\partial x}\Big(g(x) + C(x+p)\Big)\Big|_{x=\bar{s}-p} = g'(\bar{s} - p) + C'(\bar{s}) \geq 0,$$

where the inequality follows from monotonicity of g and (5.14). For the case where $\bar{s} - p \geq B$, we employ Lemma 5.4.2 and Assumption 5.2.1 to write

$$\begin{aligned}
\frac{\partial}{\partial x}\Big(g(x) + C(x+p)\Big)\Big|_{x=B} &= g'(B) + C'(B+p) \\
&\geq g'(B) - \frac{Q}{r}\mathbf{E}_W[g(W)] \\
&\geq g'(B) - \frac{\mathbf{E}_W[g(W)]}{\mathbf{E}[W]} \geq 0,
\end{aligned}$$

where the last inequality holds because $g(w) \leq wg'(B)$, for all $w \leq B$, which is a convexity result. □

Proof of Theorem 5.2.5: By Theorem 5.2.2(iii), we can characterize the optimal policy as

$$\begin{aligned}
\mu^*(s,w) &= \text{argmin} \;\; g(w-u) + C(s-u) \\
&\qquad \text{s.t. } 0 \leq u \leq \min\{s, w\}.
\end{aligned} \tag{5.28}$$

Note that the optimization problem in (5.28) is convex, because $g(\cdot)$ and hence, $C(\cdot)$ is convex (cf. Theorem 5.2.2(ii)). Using the change of variables

$$x = w - u, \quad p = s - w,$$

we can rewrite (5.28) as $\mu^*(s,w) = w - x^*(p,w)$, where

$$\begin{aligned} x^*(p,w) \;=\; & \text{argmin} \;\; g(x) + C(p+x) \\ & \text{s.t. } x \geq \max\{0, -p\} \\ & x \leq w. \end{aligned} \tag{5.29}$$

The optimization problem in (5.29) depends on both parameters p and w. We may remove the dependency on w as follows. Since $w \leq B, \bar{s} - p$, we may relax the last constraint, $x \leq w$, by replacing it with $x \leq \min\{B, \bar{s} - p\}$ The optimal solution of the relaxed problem is the same as $\phi(p)$ defined in (5.18). If $\phi(p) < w$, then the relaxed constraint is not active, and $\phi(p)$ is also the solution of (5.29). Otherwise, since we have a convex problem, the constraint $x \leq w$ must be active, which uniquely identifies the optimal solution as w. Therefore, the optimal solution of the problem in (5.29) is given by $x^*(p,w) = \min\{\phi(p), w\}$. Combining the preceding relations, we obtain

$$\mu^*(s,w) = w - \min\{\phi(s-w), w\} = \Big[w - \phi(s-w)\Big]^+.$$

The representation in (5.19) is a direct consequence of Lemmas 5.4.1 and 5.4.3. Between some break-points b_0 and b_1, the optimal solution of (5.18) can only be an interior solution, which is given by (5.20). The uniqueness of $\phi^\circ(p)$ follows from strict convexity of g. Finally, by continuous differentiability of the cost function, equation (5.20) should hold at the break-points as well. Therefore,

$$g'(b_0) + C'(b_0 + (-b_0)) = 0, \quad g'(0) + C'(0 + b_1) = 0,$$

which is equivalent to the characterizations in (5.21) and (5.22). The first inequality in (5.21) holds by Lemma 5.27 and convexity of $g(\cdot)$, and the second inequality holds by Assumption 5.2.1 and applying convexity of $g(\cdot)$ again. ■

Lemma 5.4.4. *Let $\phi(p)$ be defined as in (5.18), and assume that Assumption 5.2.1*

holds and the stage cost $g(\cdot)$ is strictly convex. Then for all $p_1 \leq p_2$,

$$-(p_2 - p_1) \leq \phi(p_2) - \phi(p_1) \leq 0. \tag{5.30}$$

Proof. We first establish the monotonicity of $\phi(p)$. Let $p_1 < p_2$. Given the structure of the kernel function in (5.19), there are multiple cases to consider, for most of which the claim is immediate using (5.19). We only present the case where $-B \leq p_1 \leq b_1$ and $b_0 \leq p_2 \leq b_1$. A necessary optimality condition at p_1 is given by

$$g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) \geq 0. \tag{5.31}$$

Similarly, for p_2, we must have

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) = 0, \tag{5.32}$$

Now, assume $\phi(p_2) > \phi(p_1)$. By convexity of $C(\cdot)$ (cf. Theorem 5.2.2(ii)) and strict convexity of $g(\cdot)$, we obtain

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) > g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) \geq 0,$$

which is a contradiction to (5.32).

For the second part of the claim, again, we should consider several cases depending on the interval to which p_1 and p_2 belong. Here, we present the case where $b_0 \leq p_1 \leq b_2$ and $b_0 \leq p_2 \leq \bar{s}$. The remaining cases are straightforward using (5.19). In this case, we have

$$g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) = 0, \tag{5.33}$$

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) \geq 0. \tag{5.34}$$

Combine the optimality conditions in (5.33) and (5.34) to get

$$g'(\phi(p_2)) + C'(p_2 + \phi(p_2)) \geq g'(\phi(p_1)) + C'(p_1 + \phi(p_1)) \tag{5.35}$$

Assume $\phi(p_2) < \phi(p_1)$; otherwise, the claim is trivial. By strict convexity of $g(\cdot)$, we have $g'(\phi(p_2)) < g'(\phi(p_1))$. Therefore by (5.35), it is true that

$$C'(p_2 + \phi(p_2)) > C'(p_1 + \phi(p_1)). \tag{5.36}$$

Now assume $\phi(p_2) - \phi(p_1) < -(p_2 - p_1)$. By rearranging the terms of this inequality and invoking the convexity of $C(\cdot)$, we get $C'(p_2 + \phi(p_2)) \leq C'(p_1 + \phi(p_1))$, which is in contradiction to (5.36). Therefore, the claim holds. □

Proof of Theorem 5.2.4: First, note that by Lemma 5.4.4, we get

$$\phi(s_2 - w) \leq \phi(s_1 - w), \quad \text{for all } w, s_1 \leq s_2$$

which implies (cf. Theorem 5.2.5)

$$\mu^*(s_2, w) = \left[w - \phi(s_2 - w)\right]^+ \geq \left[w - \phi(s_1 - w)\right]^+ = \mu^*(s_1, w).$$

Moreover, for all s and $w_1 \leq w_2$, we can use the second part of Lemma 5.4.4 to conclude

$$\phi(s - w_1) - \phi(s - w_2) \geq -(w_2 - w_1).$$

By rearranging the terms, it follows that

$$\mu^*(s, w_2) = \left[w_2 - \phi(s - w_2)\right]^+ \geq \left[w_1 - \phi(s - w_1)\right]^+ = \mu^*(s, w_1),$$

which completes the proof. ■

Chapter 6

Conclusions and Future Research Directions

6.1 Summary

We presented a new framework for studying media streaming systems in volatile environments, with focus on quality of user experience. We proposed two intuitive metrics that essentially capture the notion of *delay* from the end-user's point of view. The proposed metrics in the context of media streaming are initial buffering delay, and probability of interruption in media playback. These metrics are tractable enough to be used as a benchmark for system design.

We first characterized the optimal trade-off curves between these metrics in a single-server single-receiver configuration. These trade-offs are in analogy with the information theoretic rate-delay-reliability trade-offs for reliable communication over a noisy channel. We modeled volatility of the environment by assuming that the packet delivery from the server to the receiver is governed by a stochastic process. Under Poisson arrival process, characterized the minimum initial buffering needed to meet certain requirement in terms of interruption probability. We further generalized the results to processes with memory such as a Markov modulated Poisson process, and provided similar characterizations.

We extended the QoE trade-off characterizations to a technology-heterogeneous

multi-server system. The main challenge in multi-server systems is inefficiencies in multi-path streaming due to duplicate packet reception. This issue also significantly complicates the analysis. We proposed Network Coding as the solution to this challenge. By sending random linear combination of packets, we remove the notion of identity from packets, and hence, guarantee that no packet is redundant. Using this approach allows us to significantly simplify the flow control of multi-path streaming scenarios, and model heterogeneous multi-server systems as a single-server system.

Equipped with tools provided by network coding, we added another level of complexity to the multi-server system. We used our framework to study multi-server systems when the access cost varies across different servers. We designed various control policies to minimize the access cost while satisfying the QoE requirements in terms of buffering delay and interruption probability. We formulated the problem as a Markov decision problem with probabilistic constraints, and characterized the optimal policy by the HJB equation. In particular, we demonstrated that a simple threshold policy performs the best. The threshold policy uses the costly server if and only if the receiver's buffer is below a certain threshold. Moreover, we observed that even rare but properly timed usage of alternative access technologies may significantly improve user experience.

The tools and techniques developed in this work essentially address the transient behavior of the system. We may employ these tools in the context of other delay-sensitive applications, insurance and financial industries, and inventory control problems. In particular, we examined an analogous problem in the context of power supply networks with uncertainty in supply/demand and upward ramp constraints on both supply and storage. The uncertainty was modeled as a compound poisson arrival of energy deficit shocks. We formulated the problem of optimal control of storage for maximizing system reliability as minimization over all stationary Markovian control policies, of the infinite horizon expected discounted cost of blackouts. We showed that for a linear stage cost, a *myopic* policy which uses storage to compensate for all shocks regardless of their size is optimal. However, for strictly convex stage costs the myopic policy is not optimal. We provided a characterization of the optimal

policy strict convex stage cost functions. Our results suggest that for high ratios of the average rate of shock size to storage ramp rate, there is a critical level of storage size above which, the value of additional capacity quickly diminishes. Moreover, for all control policies, there seems to be a critical level of storage size, above which the probability of suffering large blackouts diminishes quickly.

In the following, we discuss several promising research directions with focus on media streaming applications.

6.2 Multi-user Streaming and P2P Networks

In this thesis, we discussed various scenarios of streaming a media file to a single receiver in an unreliable and volatile environment. An important extension of this problem is the problem of dynamic resource allocation in a network when streaming to multiple users. Here, allocating resources to each user can affect the quality of experience of other users. Our goal is to design resource allocation policies when taking into account such coupling effects. There are two fundamentally different scenarios to study.

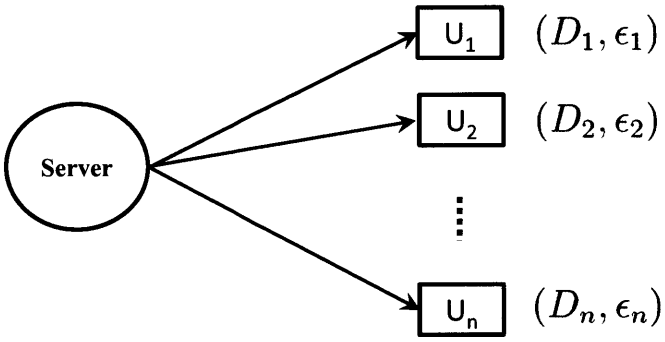

Figure 6-1: Multiple unicast streaming to multiple users with different QoE requirements.

The first case is the multiple unicast scenario. The simplest form of this setup consists of a single server streaming n different media files to n receivers (see Figure 6-1). Each user i has a QoE requirement $V_i = (D_i, \epsilon_i)$, where D_i denotes the initial buffer size of user i, and ϵ_i is his/her desired interruption probability. The total resource at

the server is limited, so that it can push at most R packets per unit of time to different users. At any point in time, the server should decide which user to serve next. The question of interest is whether all of the QoE requirements $(V_1, \ldots, V_n)$ can be *achieved* by any control policy at the server. The first step is to characterize the largest set of achievable QoE metrics given the servers limited resources. We refer to this set as the *QoE region.* Moreover, we would like to design a *QoE optimal* control policy that satisfies every user's QoE requirement for any point in the QoE region. This plan is in analogy of the queueing theory literature, where throughput optimal policies are designed to stabilize the queues for any set of admissible (achievable) arrival rates. Once this step is accomplished, we may distribute the resources among different users in a fair manner, by picking a set of QoE requirements from the QoE region that maximizes some global utility function. Owing to different channel conditions, fair streaming to different users may not be accomplished by equally dividing the available resources among the users. It is worth mentioning that we may incorporate more complicated resource constraints, e.g. interference and broadcast constraints in a multi-hop network, into our model. However, the problem formulation needs to capture the transient behavior of the system and hence, conventional network flow models are not adequate analysis tools.

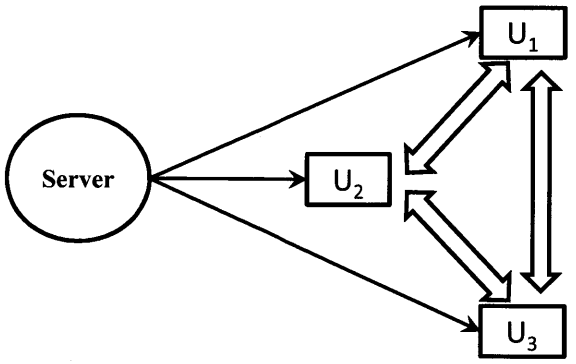

Figure 6-2: Multicast streaming with user cooperation (P2P streaming).

The second scenario of interest is Multicast/Broadcast streaming. In this setup, the same content is being streamed to multiple users. Moreover, the users can help each other in a P2P fashion (See Figure 6-2). There are multiple interesting questions

that arise in this setup. First, consider the case where users are fully cooperative. The server can decide how to allocate the resources (push packets), so that the users can achieve the largest set of QoE requirements. In this case, the users that are leading (in packet reception) due to better channel conditions, etc, should help the lagging users. Our goal is to understand whether the sever should boost the lagging users so that they can keep up, or continue helping the leading users and rely on them helping the lagging users. The solution, of course, heavily depends on system parameters such as the channel condition between the server and the users as well as the links among the users. Now, consider a pull mechanism where the users are requesting packets from the server and can act selfishly. In this scenario, a free-riding problem arises. Even though, the users should help others if they are leading in packet reception, they might have an incentive to not request the packets from the main server. This way, they will be lagging with respect to most of the other peers, and hence, they can receive most the packets from other peers at no charge. In such a situation, the delay-sensitive information may propagate throughout the network very slowly. Therefore, the same users may get very close to an interruption in playback, and decide to request the packets from the main server. But at this point, the server cannot support all of the requests. The goal is to characterize the *dynamic* behavior of the users in such a scenario.

6.3 Streaming with Imperfect Feedback

Consider the problem of streaming to a single user over an unreliable channel. One of the important underlying assumptions in our model has been the fact that the server knows which (coded or uncoded) packets have been received at any point. Therefore, it can decide whether to move to the next packet or block of packets, and hence, it always operate at maximum throughput without incurring any additional decoding delay. This can be achieved using an ARQ mechanism over a perfect feedback channel with no delay. From a more practical point of view, we need to relax this assumption. It is well-known that using random linear network coding across packets in a larger

block increases the throughput as well as decoding delay. However, when decoding delay becomes important and perfect feedback is not available, we need to code within a smaller block of packets. Clearly, there is a trade-off between the throughput and the decoding delay. We would like to address this problem in a dynamic setting. Observe that by dynamically changing the coding block size, we may operate at different points of the optimal throughput-delay trade-off curve. In our streaming framework, as the receiver's queue-length changes, the urgency for decoding packets also changes. For example, as the queue length approaches zero (close to interruption), we need the received packets to be immediately decodable, because they need to be consumed immediately by the application layer. On the other hand, if there are enough packets buffered at the receiver, the decoding delay of the current block does not play an important role, and we may increase the block size to maximize the throughput. An interesting problem is to design optimal block size selection policies that dynamically choose an appropriate set of packets to code and transmit. Such policies need to be robust againts imperfect queue-length information received from the end-user.

Bibliography

[1] C. Labovitz, D. McPherson, and S. Iekel-Johnson. 2009 Internet Observatory report. In *NANOG-47*, October 2009.

[2] Cisco visual networking index: Forecast and methodology, 2009-2014. Cisco, 2010.

[3] R. Laroia. Future of Wireless? The Proximate Internet. In *Proc. of the Second International Conference on Communication Systems and Networks (COMSNETS)*, Bangalore, India, January 2010.

[4] Knocking. http://knockinglive.com, 2010.

[5] R.C. Chen. Constrained stochastic control with probabilistic criteria and search optimization. In *Proc. 43rd IEEE Conference on Decision and Control*, December 2004.

[6] M. J. Neely. Super-fast delay tradeoffs for utility optimal fair scheduling in wireless networks. *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Nonlinear Optimization of Communication Systems*, 24(8):1489–1501, 2006.

[7] M. J. Neely. Delay-based network utility maximization. In *Proceedings of INFOCOM*, San Diego, CA, March 2010.

[8] S. Lu, V. Bharghavan, and R. Srikant. Fair scheduling in wireless packet networks. *IEEE/ACM Transactions on Networking (TON)*, 7(4):473–489, 1999.

[9] R.A. Berry and R.G. Gallager. Communication over fading channels with delay constraints. *IEEE Transactions on Information Theory*, 48(5):1135–1149, 2003.

[10] I. Hou and P. R. Kumar. Scheduling heterogeneous real-time traffic over fading wireless channels. In *Proceedings of INFOCOM*, San Diego, CA, March 2010.

[11] Y. Zhou, D. Chiu, and J. Lui. A simple model for analyzing P2P streaming protocols. In *Proc. IEEE ICNP 2007*.

[12] T. Bonald, L. Massoulié, F. Mathieu, D. Perino, and A. Twigg. Epidemic live streaming: optimal performance trade-offs. *SIGMETRICS Perform. Eval. Rev.*, 36(1):325–336, 2008.

[13] Bridge Q. Zhao, John C.S. Lui, and Dah-Ming Chiu. Exploring the optimal chunk selection policy for data-driven P2P streaming systems. In *The 9th International Conference on Peer-to-Peer Computing*, 2009.

[14] L. Ying, R. Srikant, and S. Shakkottai. The Asymptotic Behavior of Minimum Buffer Size Requirements in Large P2P Streaming Networks. In *Proc. of the Information Theory and Applications Workshop*, San Diego, CA, 2010.

[15] V. N. Padmanabhan, H. J. Wang, P. A. Chou, and K. Sripanidkulchai. Distributing Streaming Media Content Using Cooperative Networking. In *Proceedings of The 12th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV '02)*, Miami, FL, May 2002.

[16] E. Setton and J. Apostolopoulos. Towards Quality of Service for Peer-to-Peer Video Multicast. In *Proc. of IEEE International Conference on Image Processing (ICIP)*, San Antonio, TX, September 2007.

[17] S. Liu, R. Zhang-Shen, W. Jiang, J. Rexford, and M. Chiang. Performance bounds for peer-assisted live streaming. In *Proc. ACM SIGMETRICS*, June 2008.

[18] N. Magharei, R. Rejaie, and Y. Guo. Mesh or multiple-tree: A comparative study of live P2P streaming approaches. In *Proc. IEEE INFOCOM*, Anchorage, AK, May 2007.

[19] B. Cohen. Incentives to build robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, June 2003.

[20] X. Zhang, J. Liu, B. Li, and T.-S. P. Yum. Coolstreaming/donet: A data-driven overlay network for efficient live media streaming. In *Proc. IEEE INFOCOM*, Miami, FL, March 2005.

[21] PPLive. http://www.pplive.com/, 2009.

[22] QQLive. http://www.qqlive.com/, 2009.

[23] TVAnts. http://www.tvants.com/, 2009.

[24] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger. A random linear network coding approach to multicast. *IEEE Transactions on Information Theory*, 52:4413–4430, 2006.

[25] S. Acedanski, S. Deb, M. Médard, and R. Koetter. How good is random linear coding based distributed networked storage. In *Proc. NetCod*, 2005.

[26] C. Gkantsidis, J. Miller, and P. Rodriguez. Comprehensive view of a live network coding P2P system. In *Proc. ACM SIGCOMM*, 2006.

[27] M. Wang and B. Li. R^2: Random push with random network coding in live peer-to-peer streaming. *IEEE JSAC, Special Issue on Advances in Peer-to-Peer Streaming Systems*, 25:1655–1666, 2007.

[28] D. Kumar, E. Altman, and J-M. Kelif. Globally Optimal User-Network Association in an 802.11 WLAN and 3G UMTS Hybrid Cell. In *Proc. of the 20th International Teletraffic Congress (ITC-20)*, Ottawa, 2007.

[29] S. Asmuseen. *Ruin Probabilities.* World Scientific Publishing Company, 2000.

[30] J. M. Reinhard. On a class of semi-Markov risk models obtained as classical risk models in a Markovian environment. *ASTIN Bull*, 14(1):23–43, 1984.

[31] E. Altman. *Constrained Markov Decision Processes*. Chapman and Hall, 1999.

[32] A.B. Piunovskiy. *Optimal Control of Random Sequences in Problems with Constraints*. Kluwer Academic Publishers, 1997.

[33] A.B. Piunovskiy. Controlled random sequences: the convex analytic approach and constrained problems. *Russian Mathematical Surveys*, 53:1233–1293, 1998.

[34] E.A. Feinberg and A. Shwartz. Constrained Markov decision models with discounted rewards. *Mathematics of Operations Research*, 20:302–320, 1995.

[35] G.L. Blankenship R.C. Chen. Dynamic programming equations for discounted constrained stochastic control. *IEEE Transactions on Automatic Control*, 49:699–709, 2004.

[36] A.B. Piunovskiy and X. Mao. Constrained Markov decision processes: the dynamic programming approach. *Operations Research Letters*, 27:119–126, 2000.

[37] R. Brockett. Stochastic control. Lecture Notes, Harvard University, Cambridge, MA, 2009.

[38] A. Faghih, M. Roozbehani, and M. A. Dahleh. Optimal utilization of storage and the induced price elasticity of demand in the presence of ramp constraints. *Decision and Control (CDC), 2011 50th IEEE Conference on*, To Appear, 2011.

[39] M. Chen, I.-K. Cho, and S.P. Meyn. Reliability by design in a distributed power transmission network. *Automatica*, 42:1267–1281, August 2006. (invited).

[40] I.-K. Cho and S. P. Meyn. A dynamic newsboy model for optimal reserve management in electricity markets. Submitted for publication, SIAM J. Control and Optimization., 2009.

[41] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris. *Fundamentals of queueing theory*. John Wiley & Sons, Hoboken, NJ, 4th edition, 2009.

[42] Sean Meyn. *Control Techniques for Complex Networks.* Cambridge University Press, New York, NY, USA, 1st edition, 2007.

[43] R. G. Gallager. *Information Theory and Reliable Communication.* Wiley, 1968.

[44] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 1997.

[45] P. W. Glynn. Upper bounds on Poisson tail probabilities. *Operations Research Letters*, 6(1):9–14, 1987.

[46] B. L. Fox and P. W. Glynn. Computing Poisson probabilities. *Communications of the ACM*, 31(4):440–445, 1988.

[47] D. Wischik, M. Handley, and C. Raiciu. Control of multipath tcp and optimization of multipath routing in the internet. *Network Control and Optimization*, 5894:204–218, 2009.

[48] D. Wischik, M. Handley, and C. Raiciu. Control of multipath tcp and optimization of multipath routing in the internet. *Network Control and Optimization*, 5894:204–218, 2009.

[49] A. Ford, C. Raiciu, M. Handley, S. Barre, and J. Iyengar. Architectural guidelines for multipath tcp development. July 2011. draft-ietf-mptcp-architecture-05.

[50] Jay Kumar Sundararajan, Szymon Jakubczak, Muriel Médard, Michael Mitzenmacher, and João Barros. Interfacing network coding with tcp: an implementation. *CoRR*, abs/0908.1564, 2009.

[51] T. Kurtz. Strong approximation theorems for density dependent markov chains. *Stochastic Processes and Their Applications*, 6:223–240, 1978.

[52] D. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1,2. Athena Scientific, 3rd edition, 2007.

[53] L.M. Benveniste and J.A. Scheinkman. On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47(3), 2010.

[54] W.H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer, 2nd edition, 2005.

[55] R. J. Elliott. *Stochastic Calculus and Applications.* SpringerVerlag, 1982.