The Dissertation Committee for Vinay Joseph
certifies that this is the approved version of the following dissertation:

Mean-Variability-Fairness Tradeoffs in Resource Allocation with Applications to Video Delivery

Committee:

Gustavo de Veciana, Supervisor

Al Bovik

Constantine Caramanis

David Morton

Sanjay Shakkottai

Mean-Variability-Fairness Tradeoffs in Resource Allocation with Applications to Video Delivery

by

Vinay Joseph, B.Tech., M.E.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2013

This thesis is dedicated to my parents.

Acknowledgments

I thank God for all the blessings gifted to me and, in particular, for all the opportunities to pursue education. I am especially grateful for my family, the most important blessing of all. I am greatly indebted to my parents, brother, grandparents and the rest of my family for their love, sacrifice and constant support.

I am fortunate to have had an exceptional researcher/teacher/mentor like Prof. Gustavo de Veciana as my PhD advisor. I firstly thank him for the generous amount of time and effort that he has spent over the past four years on numerous discussions and documentation related to this thesis. These discussions, filled with valuable insightful comments from my advisor, form the core of this thesis. I am especially grateful for his tremendous patience during the numerous times we got stuck in our work, and for all the flexibility provided in steering the direction of my research. The patience and flexibility played a vital role in turning my PhD work into one of the most enjoyable and intellectually satisfying endeavors of my life.

I thank Dr. Ari Arapostathis for all the helpful discussions and ideas that have greatly improved this thesis. I also thank my Master's program advisors Prof. Vinod Sharma and Prof. Utpal Mukherji for providing me a healthy first course of research. I thank Sem Borst and Marty Rieman for a great summer at Bell Labs. I also thank members of my PhD committee Profs. Sanjay Shakkottai, David Morton, Constantine Caramanis and Al Bovik for their helpful comments.

I thank Zheng Lu for numerous helpful discussions and help with various simulations. I also thank Xiaoqing Zhu, Chao Chen and Sarabjot Singh for all the discussions that have helped us to improve various aspects of our work.

I also thank Intel/Cisco and NSF for supporting the work in this thesis through the VAWN program and the NSF Grant CNS-0917067 respectively.

I am grateful to all my teachers for their help in learning many of the ideas that have influenced this thesis. I also thank all my friends and colleagues for their help and company during these special four years of my life.

VINAY JOSEPH

The University of Texas at Austin
August 2013

Mean-Variability-Fairness Tradeoffs in Resource Allocation with Applications to Video Delivery

Publication No. ____________________

Vinay Joseph, Ph.D.
The University of Texas at Austin, 2013

Supervisor: Gustavo de Veciana

Network Utility Maximization (NUM) provides a key conceptual framework to study reward allocation amongst a collection of users/entities in disciplines as diverse as economics, law and engineering. However when the available resources and/or users' utilities vary over time, reward allocations will tend to vary, which in turn may have a detrimental impact on the users' overall satisfaction or quality of experience. In this thesis, we introduce a generalization of the NUM framework which incorporates the detrimental impact of temporal variability in a user's allocated rewards and explicitly incorporates Mean-Variability-Fairness tradeoffs, i.e., tradeoffs amongst the mean and variability in users' reward allocations, as well as fairness across users. We propose a simple online algorithm to realize these trade-

offs, which, under stationary ergodic assumptions, is shown to be asymptotically optimal, i.e., achieves a long term performance equal to that of an offline algorithm with knowledge of the future variability in the system. This substantially extends work on NUM to an interesting class of relevant problems where users/entities are sensitive to temporal variability in their service or allocated rewards.

We extend the theoretical framework and tools developed for realizing Mean-Variability-Fairness tradeoffs to develop a simple online algorithm to solve the problem of optimizing video delivery in networks. The tremendous increase in mobile video traffic projected for the future along with insufficiency of available wireless network capacity makes this one of the most important networking problems today. Specifically, we consider a network supporting video clients streaming stored video, and focus on the problem of jointly optimizing network resource allocation and video clients' video quality adaptation. Our objective is to fairly maximize video clients' video Quality of Experience (QoE) realizing Mean-Variability-Fairness tradeoffs, incorporating client preferences on rebuffering time and the cost of video delivery. We present a simple asymptotically optimal online algorithm NOVA (Network Optimization for Video Adaptation) to solve the problem. Our algorithm uses minimal communication, 'distributes' the tasks of network resource allocation to a centralized network controller, and video clients' video quality adaptation to the respective video clients. Further, the quality adaptation is also optimal for standalone video clients, and is an asynchronous algorithm well suited for use in the Dynamic Adaptive Streaming over HTTP (DASH) framework.

We also extend NOVA for use with more general video QoE models, and study NOVA accounting for practical considerations like time varying number of video clients, sharing with other types of traffic, performance under legacy resource allocation policies, videos with variable sized segments etc.

Contents

Chapter 1

Introduction

Network Utility Maximization (NUM) is a key conceptual framework to study reward allocation among a collection of users/entities across disciplines as diverse as economics, law and engineering. In network engineering, the NUM framework has served as a particularly insightful setting to study (reverse engineer) how the Internet's congestion control protocols allocate bandwidth, how to devise schedulers for wireless systems with time varying channel capacities, and also motivated the development of distributed mechanisms to maximize network utility in diverse settings including communication networks and the smart grid, while incorporating new relevant constraints, on energy, power, storage, power control, stability, etc.

When the available resources/rewards and/or users' utilities vary over time, reward allocations amongst users will tend to vary, which in turn may have a detrimental impact on the users' utility or perceived service quality. Indeed temporal variability in utility, service, rewards or associated prices are particularly problematic when humans are the eventual recipients of the allocations. Humans typically view temporal variability negatively, as a sign of an unreliable service, network or market instability. Broadly speaking, temporal variability, when viewed through human's cognitive and behavioral responses, leads to a degraded Quality of Experience (QoE). This in turn can lead users to make decisions, e.g., change provider, act upon perceived market instabilities, etc., which can have serious implications on businesses and engineered systems, or economic markets. For problems involving resource allocation in networks, [9] argues that predictable or consistent service is essential and even points out that it may be appropriate to intentionally lower the quality delivered to the user if that level is sustainable.

For a user viewing a video stream, variations in video quality over time have a detrimental impact on the user's QoE, see e.g., [59, 28, 40]. Indeed [59] suggested that variations in quality can result in a QoE that is worse than that of a constant quality video with lower average quality. Furthermore, [59] proposed a metric for QoE given below which penalizes standard deviation of quality over time:

$$\text{Mean Quality} - \kappa\sqrt{\text{Temporal Variance in Quality}}$$

where κ is an appropriately chosen positive constant. [19] and [53] argue that less variability in the service processes can improve customer satisfaction by studying data for large retail banks and major airlines respectively. Aversion towards temporal variability is not just restricted to human behavior, for instance, see [38] for a discussion of the impact of temporal variability in nectar reward on foraging behavior of bees. Also, variability in resource allocation in networks can lead to burstiness which can degrade network performance (see [11, 41]). These examples illustrate the need for extending the NUM framework to incorporate the impact of variability.

In Chapter 2, we develop a generalized NUM framework which explicitly incorporates the detrimental impact of temporal variability in a user's allocated rewards. We use the term rewards as a proxy for the resulting utility of, or any other quantity associated with, resource allocations to users/entities in a system. For instance, in wireless network serving video users, resource allocation concerns decisions about allocation of resources like bandwidth, power etc., and the resulting video quality corresponds to the reward. Our goal is to explicitly tackle the task of incorporating tradeoffs amongst the mean and variability in users' rewards. Thus, for example, in a variance-sensitive NUM setting, it may make sense to reduce a user's mean reward so as to reduce his/her variability. There are many ways in which temporal variations can be accounted for, and which, in fact, present distinct technical challenges. In this thesis, we shall take a simple elegant approach to the problem which serves to address systems where tradeoffs amongst the mean and variability over time need to be made rather than systems where the desired mean (or target) is known (as in minimum variance control, see [4]), or where the issue at hand is minimization of the variance of a cumulative reward at the end of a given (e.g., investment) period.

Chapter 2 contains one of the major contributions of this thesis: the development of a *simple online* algorithm, Adaptive Variability-aware Reward allocation (AVR), to solve problems falling in the generalized NUM framework. Under stationary ergodic assumptions, AVR is shown to be asymptotically *optimal*, i.e., achieves a long term performance equal to that of an optimal omniscient offline algorithm.

In Chapter 3, we extend the theoretical framework and tools developed in Chapter 2 to solve the problem of optimizing stored video delivery in networks. In particular, we study an *asynchronous* extension of the reward allocation framework studied in Chapter 2. The reward allocation framework studied in Chapter 2 is synchronous in the sense that the decisions concerning reward allocations and the associated resource allocation are made in a synchronous manner. In Chapters3, we study a reward allocation framework in which reward allocation decisions are made in an asynchronous manner, and which allows 'buffering'/delaying of eventual rewards. This feature is particularly useful in Chapter 3 where we consider a setting in which users stream/download long videos stored at video servers. A long video file can be viewed as a concatenation of several short files called seg-

Table 1.1: A comparison of frameworks in Chapters 2 and 3

Framework in	Synchronous resource allocation?	Synchronous reward allocation?	'Buffered' reward allocation?	Resource-Reward coupling
Chapter 2	Yes	Yes	No	Instantaneous
Chapter 3	Yes	No	Yes	Averaged

ments, a user downloads the video by downloading the segments sequentially, and the reward allocation decisions in this setting correspond to decisions about the segments' quality. Although decisions concerning the underlying resource allocation (e.g., bandwidth allocation) are made in a synchronous manner, the reward allocation decision associated with a segment, i.e., the decision concerning the segment's quality, is made only after the completion of download of the previous segment. Apart from the asynchronous nature of the framework, a key distinguishing feature of the two frameworks is the fact that, the reward allocation decisions and resource allocation decisions considered in Chapter 3 are coupled through constraints that only account for *averages* associated with these decisions, whereas there is an *instantaneous* coupling between reward allocation decisions and resource allocation decisions in the setting considered in Chapter 2. In Table 1.1, we have summarized some of the key features of the frameworks studied in Chapters 2 and 3.

In Chapter 3, we use the asynchronous reward allocation framework to model a network supporting video clients streaming stored video, and focus on the problem of jointly optimizing network resource allocation and video clients' video quality adaptation. Projections of tremendous increase in mobile video traffic and insufficiency of available wireless network capacity highlights the importance of this problem. In Chapter 3, we develop a simple online algorithm NOVA to solve this problem, and also establish its asymptotic optimality.

One of our main objectives in the development of NOVA was to ensure that it could be used with current and future practical systems, for e.g., the quality adaptation proposed in NOVA is well suited for Dynamic Adaptive Streaming over HTTP (DASH) framework. This is the state of the art framework being proposed for stored video delivery (and possibly for real time streaming applications). Hence, we go beyond the theoretical analysis of NOVA in Chapter 3, and study the performance of NOVA taking several practical considerations into account like time varying numbers of video clients, sharing with (and presence of) other types of traffic, performance under legacy resource allocation policies etc. In Chapter 4, we also study the performance of NOVA using simulations under a variety of settings using real world data.

1.1 Key contributions of this thesis

Below, we summarize the key contributions of this thesis.

I.a We develop a generalized NUM framework which explicitly accounts for Mean-Variability-Fairness tradeoffs associated with users' reward allocation.

I.b We develop a simple asymptotically optimal online algorithm AVR (Adaptive Variability-aware Reward allocation) to solve problems falling in this framework.

II.a We propose a general optimization framework for stored video delivery optimization, that factors heterogeneity in client preferences, QoE models (that account for Mean-Variability tradeoffs in quality), capacity and video content.

II.b We develop a *simple online* algorithm NOVA (Network Optimization for Video Adaptation) to solve the video delivery optimization problem. Key features of NOVA are listed below:

1. *Optimality:* We establish a strong asymptotic optimality result for NOVA which roughly guarantees that NOVA performs as well as the optimal offline scheme which is omniscient, i.e., knows everything about the evolution of channel and video ahead of time.

2. *Simple and Online*: NOVA only utilizes current information, and is computationally light.

3. *Distributed*: NOVA uses minimal signaling, and can be implemented in a distributed manner.

4. NOVA is *asynchronous* and requires almost no statistical information about the system

5. *Optimal Adaptation:* The adaptation proposed in NOVA is independently optimal, and the optimality properties of the adaptation component of NOVA is 'insensitive' to the resource allocation, i.e., does not depend on detailed characteristics (for e.g., the specific resource allocation algorithm, time scale of operation etc) of the latter. Further, the adaptation proposed in NOVA is entirely client driven and (can be used for and) is also optimal for standalone video clients.

6. *Suited for current practical systems:*

   (a) *Suited for DASH:* The adaptation proposed in NOVA is suited for DASH framework as it entirely client driven, and can be carried out in an asynchronous manner.

   (b) The resource allocation proposed in NOVA requires simple modification of legacy schedulers like proportionally fair schedulers, and can be extended for use in the presence of data users.

1.2 Organization of this thesis

In Chapter 2, we develop the generalized NUM framework to realize optimal Mean-Variability-Fairness tradeoffs, develop the algorithm AVR to solve problems falling in this framework, and establish the optimality of AVR. In Chapter 3, we study the problem of stored video delivery optimization, present the algorithm NOVA to solve the problem, and establish the optimality of NOVA. In Chapter 4, we discuss the performance of NOVA taking several practical considerations into account and study the performance of NOVA using simulations. We conclude the thesis in Chapter 5 with a discussion about some future directions and open problems.

Chapter 2

Mean-Variability-Fairness Tradeoffs in Resource Allocation

2.1 Introduction

Network Utility Maximization (NUM) is a key conceptual framework to study (fair) reward allocation among a collection of users/entities across disciplines as diverse as economics, law and engineering. For example, [43] introduces NUM for realizing fair allocations of a *fixed* amount of water c to N farms. The amount of water w_i allocated to the ith farm is a resource which yields a reward $r_i = f_i(w_i)$ to the ith farm. Here, f_i is a concave function mapping allocated water (resource) to yield (reward), and these can differ across farms. The allocation maximizing $\sum_{1 \le i \le N} r_i$ is a reward (utility) maximizing solution to the problem. Fairness can be imposed on the allocation by changing the objective of the problem to $\sum_{1 \le i \le N} U(r_i)$ for an appropriately chosen concave function U. Now, suppose that we have to make allocation decisions periodically to respond to time varying water availability $(c_t)_{t \in \mathbb{N}}$ and utility functions $(f_{i,t})_t$. Then, subject to the time varying constraints, one could maximize (see e.g., [49], [30])

$$\sum_{1 \le i \le N} U(\overline{r}_i) \tag{2.1}$$

to obtain a resource allocation scheme which is fair in the delivery of time average rewards $\overline{\mathbf{r}} = (\overline{r}_i)_{i \in \mathcal{N}}$.

In network engineering, the NUM framework has served as a particularly insightful setting to study (reverse engineer) how the Internet's congestion control protocols allocate bandwidth, how to devise schedulers for wireless systems with time varying channel capacities, and also motivated the development of distributed mechanisms to maximize network utility in diverse settings including communication networks and the smart grid, while incorporating new relevant constraints, on energy, power, storage, power control, stability, etc.

When the available resources/rewards and/or users' utilities vary over time, reward allocations amongst users will tend to vary, which in turn may have a detrimental impact on the users' utility or perceived service quality. In fact, temporal variability in farm water availability can have a negative impact on crop yield (see [47]). This motivates modifications of formulations with objectives such as the one in (2.1) to account for this impact.

Indeed temporal variability in utility, service, rewards or associated prices are particularly problematic when humans are the eventual recipients of the allocations. Humans typically view temporal variability negatively, as a sign of an unreliable service, network or market instability. Broadly speaking, temporal variability, when viewed through human's cognitive and behavioral responses, leads to a degraded Quality of Experience (QoE). This in turn can lead users to make decisions, e.g., change provider, act upon perceived market instabilities, etc., which can have serious implications on businesses and engineered systems, or economic markets. For problems involving resource allocation in networks, [9] argues that predictable or consistent service is essential and even points out that it may be appropriate to intentionally lower the quality delivered to the user if that level is sustainable.

For a user viewing a video stream, variations in video quality over time have a detrimental impact on the user's QoE, see e.g., [59, 28, 40]. Indeed [59] suggested that variations in quality can result in a QoE that is worse than that of a constant quality video with lower average quality. Furthermore, [59] proposed a metric for QoE given below which penalizes standard deviation of quality over time:

$$\text{Mean Quality} - \kappa\sqrt{\text{Temporal Variance in Quality}}$$

where κ is an appropriately chosen positive constant. [19] and [53] argue that less variability in the service processes can improve customer satisfaction by studying data for large retail banks and major airlines respectively. Aversion towards temporal variability is not just restricted to human behavior, for instance, see [38] for a discussion of the impact of temporal variability in nectar reward on foraging behavior of bees. Also, variability in resource allocation in networks can lead to burstiness which can degrade network performance (see [11, 41]). These examples illustrate the need for extending the NUM framework to incorporate the impact of variability.

This chapter introduces a generalized NUM framework which explicitly incorporates the detrimental impact of temporal variability in a user's allocated rewards. We use the term rewards as a proxy for the resulting utility of, or any other quantity associated with, allocations to users/entities in a system. Our goal is to explicitly tackle the task of incorporating tradeoffs amongst the mean and variability in users' rewards. Thus, for example, in a variance-sensitive NUM setting, it may make sense to reduce a user's mean reward so as to reduce his/her variability. As will be discussed in the sequel, there are many ways in which temporal variations can be accounted for, and which, in fact, present distinct technical challenges. In this chapter, we shall take a simple elegant approach to the problem which serves to address systems where tradeoffs amongst the mean and variability over time need to be made rather than systems where the desired mean (or target) is known (as in minimum variance control, see [4]), or where the issue at hand is minimization of the variance of a cumulative reward at the end of a given (e.g., investment) period.

To better describe the characteristics of the problem we introduce some preliminary notation. We shall consider a network shared by a set $\mathcal{N}$ of users (or other entities) where $N{:=}|\mathcal{N}|$ denotes the number of users in the system. Throughout the chapter, we distinguish between random variables (and random functions) and their realizations by using upper case letters for the former and lower case for the latter. Let $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, real numbers and nonnegative real numbers respectively. We use bold letters to denote vectors, e.g., $\mathbf{a} = (a_i)_{i\in\mathcal{N}}$. Given a collection of T objects $(b(t))_{1\leq t\leq T}$ or a sequence $(b(t))_{t\in\mathbb{N}}$, we let $(b)_{1:T}$ denote the finite length sequence $(b(t))_{1\leq t\leq T}$ (in the space associated with the objects of the sequence). For example, consider a sequence $(\mathbf{b}(t))_{t\in\mathbb{N}}$ where each element is a vector. Then $(\mathbf{b})_{1:T}$ denotes the T length sequence containing the first T vectors of the sequence $(\mathbf{b})_{1:T}$, and $(b_i)_{1:T}$ denotes the sequence containing the ith component of the first T vectors. For any function U on $\mathbb{R}$, let U' denote its derivative.

Definition 2.1. For any (infinite length) sequence of real numbers $(a(t))_{t\in\mathbb{N}}$, let

$$\begin{aligned}
m^T(a) &:= \frac{1}{T}\sum_{t=1}^{T} a(t), \\
\mathrm{Var}^T(a) &:= \frac{1}{T}\sum_{t=1}^{T} \left(a(t) - m^T(a)\right)^2, \\
e_i^T(a) &:= m^T(a) - U_i^V\left(\mathrm{Var}^T(a)\right),
\end{aligned}$$

i.e., $m^T(a)$ and $\mathrm{Var}^T(a)$ denote empirical mean and variance. Note that the argument a used in the functions $m^T(a)$, $\mathrm{Var}^T(a)$ and $e_i^T(a)$ stands for the associated sequence $(a(t))_{t\in\mathbb{N}}$. We will also (abusing notation) use the above operators on any finite length sequence $(a)_{1:T} \in \mathbb{R}^T$ of real numbers.

Let $r_i(t)$ represent the reward allocated to user i at time t. Then $\mathbf{r}(t) = (r_i(t))_{i\in\mathcal{N}}$ is the vector of rewards to users $\mathcal{N}$ at time t, and $(\mathbf{r})_{1:T}$ represents sequence of vector rewards allocated over time slots $t = 1, \ldots, T$. We assume that reward allocations are subject to time varying network constraints,

$$c_t(\mathbf{r}(t)) \leq 0 \;\; \text{for} \;\; t = 1, \ldots, T,$$

where each $c_t : \mathbb{R}^N \to \mathbb{R}$ is a convex function, thus implicitly defining a convex set of feasible reward allocations. To formally capture the impact of the time-varying rewards on users' QoE consider the following *offline* convex optimization problem OPT(T):

$$\max_{(\mathbf{r})_{1:T}} \sum_{i\in\mathcal{N}} U_i^E \left(\overbrace{\underbrace{m^T(r_i)}_{\text{Mean Reward}} \quad - \quad \underbrace{U_i^V\left(\mathrm{Var}^T(r_i)\right)}_{\text{Penalty for Variability}}}^{\text{User } i\text{'s QoE}} \right),$$

$$\text{subject to} \qquad c_t(\mathbf{r}(t)) \leq 0, \;\; \mathbf{r}(t) \geq \mathbf{0} \;\; \forall\, t \in \{1, ..., T\}.$$

We refer to OPT(T) as an offline optimization because time-varying time constraints $(c_t)_{1:T}$ are assumed to be known. Here, we introduce increasing functions $\left(U_i^E, U_i^V\right)_{i\in\mathcal{N}}$ such that the above optimization problem is convex. For user i, the argument of the function U_i^E is our proxy for the user's QoE. Thus, the desired fairness in the allocation of QoE across the users can be imposed by appropriately choosing $\left(U_i^E\right)_{i\in\mathcal{N}}$. Note that the first term $m^T(r_i)$ in user i's QoE is the user's mean reward allocation, whereas the presence of the empirical variance function $\mathrm{Var}^T(r_i)$ in the second term penalizes temporal variability in a reward allocation.

Further, flexibility in picking $\left(U_i^V\right)_{i\in\mathcal{N}}$ allows for several different ways to penalize such variability. Indeed, one can in principle have a variability penalty that is convex or concave in variance. Hence, the formulation OPT(T) allows us to realize tradeoffs among mean, fairness and variability associated with the reward allocation by appropriately choosing the functions $\left(U_i^E, U_i^V\right)_{i\in\mathcal{N}}$.

2.1.1 Main contributions

The main contribution of this chapter is the development of an *online* algorithm, Adaptive Variability-aware Reward allocation (AVR), which asymptotically solves OPT(T). The algorithm requires almost no statistical information about the system, and its characteristics are as follows:

(i) in each time slot, c_t is revealed, and AVR *greedily* allocates rewards by solving the following optimization problem OPT-ONLINE:

$$\begin{aligned} &\max_{\mathbf{r}} \sum_{i\in\mathcal{N}} \left(U_i^E\right)' (e_i(t)) \left(r_i - \left(U_i^V\right)' (v_i(t)) \left(r_i - m_i(t)\right)^2\right) \\ &\text{subject to } c_t(\mathbf{r}) \le 0, \quad \mathbf{r} \ge 0, \end{aligned}$$

where $e_i(t) = m_i(t) - U_i^V\left(v_i(t)\right)$ for each $i \in \mathcal{N}$ is an estimate of the user's QoE based on estimated means and variances $\mathbf{m}(t)$ and $\mathbf{v}(t)$; and,

(ii) it updates (vector) parameters $\mathbf{m}(t)$ and $\boldsymbol{v}(t)$ to keep track of the mean and variance of the reward allocations under AVR.

Under stationary ergodic assumptions for time-varying constraints, we show that our *online* algorithm AVR is asymptotically optimal, i.e., achieves a performance equal to that of the *offline* optimization OPT(T) introduced earlier as $T \to \infty$. This is a strong optimality result, which at first sight may be surprising due to the variability penalty on rewards and the time varying nature of the constraints $(c_t)_{t\in\mathbb{N}}$. The key idea is to keep online estimates for the relevant quantities associated with users' reward allocations, e.g., the mean and variance which over time are shown to converge. This in turn eventually enables our greedy online policy to produce reward allocations corresponding to the optimal stationary policy. Proving this result is somewhat challenging as it requires showing that the estimates based on reward allocations produced by our online policy, AVR, (which itself depends on the estimated quantities), will converge to the desired values. To our knowledge this is the first attempt to generalize the NUM framework in this direction. We contrast our problem formulation and approach to past work in addressing 'variability' minimization, risk-sensitive control and other MDP based frameworks in the next subsection.

2.1.2 Related work

Network Utility Maximization (NUM) is a well studied approach used for reward allocation amongst a collection of users/entities. The work in [43] provides a network-centric overview of NUM. All the work on NUM including several major extensions (for e.g., [27], [49], [48], [37] etc.) has ignored the impact of variability in reward allocation. Our work [24] is to our knowledge the first to tackle NUM incorporating the impact of variability explicitly. In particular, we addressed a special case of the problem studied in this chapter that only allows for linear functions $\left(U_i^E, U_i^V\right)_{i\in\mathcal{N}}$, and an asymptotically optimal online reward allocation algorithm for a wireless network supporting video streaming users is proposed. The algorithm proposed and analyzed in this chapter is a generalization of gradient based algorithms studied in [2], [30] and [49]. Our approach for proving asymptotic optimality generalizes those in [49] and [25]. In [49], the focus is on objectives such as (2.1), but does not allow for the addition of penalty terms on temporal variance in the objective. By contrast with this chapter, the approaches in [24] and [25] rely on the use of results on sensitivity analysis of optimization problems, and only allows for linear $\left(U_i^E\right)_{i\in\mathcal{N}}$

and concave $\left(U_i^V\right)_{i\in\mathcal{N}}$.

Adding a temporal variance term in the cost takes the objective out of the basic dynamic programming setting (even when $\left(U_i^E, U_i^V\right)_{i\in\mathcal{N}}$ are linear) as the overall cost is not decomposable over time, i.e., can not be written as a sum of costs each depending only on the allocation at that time- this is what makes sensitivity to variability challenging. For risk sensitive decision making, MDP based approaches aimed at realizing optimal tradeoffs between mean and temporal variance in reward/cost were proposed in [18] and [45]. While they consider a more general setting than ours where actions can even affect future feasible reward allocations, e.g., may affect the process $(C_t)_{t\in\mathbb{N}}$ itself, the approaches proposed in these works suffer from the curse of dimensionality as they require solving large optimization problems. For instance, the work of [18] involves solving a quadratic program in the (typically large) space of state-action pairs. Note that these works on risk sensitive decision making are different from those focusing on the variance of the *cumulative* cost/reward such as the one in [33].

Variability or perceived variability can be measured in many different ways, and temporal variance considered in this chapter is one of them. One could also 'reduce variability' using a minimum variance controller (see [4]) where we have certain target reward values fixed ahead of time and big fluctuations from these targets are undesirable. Note however that in using this approach, we have to fix our targets ahead of time, and thus lose the ability to realize tradeoffs between the mean and variability in reward allocation. One could also measure variability using switching costs like in [31], which consider the problem of achieving tradeoffs between average cost and time average switching cost associated with data center operation, and proposes algorithms with good performance guarantees for adversarial scenarios. The decision regarding how to penalize variability is ultimately dependent on the application setting under consideration. We summarize the key points of the discussion

Table 2.1: Realizing Mean-Variability-Tradeoffs: Related Work

	Allows variability penalties?	Strong optimality guarantees?	Simple?	Online?	Need system statistics?
Our Work	Yes	Yes	Yes	Yes	No
[2, 30, 49]	No	Yes	Yes	Yes	No
[18, 45]	Yes	No	No	No	Yes

about related work in Table 2.1.

2.1.3 Organization of the chapter

Section 2.2 introduces the system model and assumptions. Section 2.3 presents and studies the offline formulation for optimal variance sensitive joint reward allocation OPT(T). Section 2.4 formally introduces our online algorithm AVR and presents our key convergence result which is used to prove asymptotic optimality of AVR. Section 2.5 is devoted to the proof of AVR's convergence and Section 2.6 presents simulation results exhibiting additional performance characteristics of AVR. We conclude the chapter with Section 2.7.

2.2 System model

We consider a slotted system where time slots are indexed by $t \in \mathbb{N}$, and the system serves a fixed set of users $\mathcal{N}$ and let $N := |\mathcal{N}|$.

We assume that rewards are allocated subject to time varying constraints. The reward allocation $\mathbf{r}(t) \in \mathbb{R}_+^N$ in time slot t is constrained to satisfy the following inequality

$$c_t(\mathbf{r}(t)) \leq 0,$$

where c_t denotes the realization of a randomly selected function C_t from a finite set $\mathcal{C}$ of real valued maps on $\mathbb{R}_+^N$. We model the reward constraints $(C_t)_{t\in\mathbb{N}}$ as a random process where each C_t can be interchangeably viewed as a random function or an associated index for such a function which is selected from a finite set $\mathcal{C}$. We make the following assumptions on these constraints:

Assumptions C1-C3 (Time varying constraints on rewards)

C.1 $(C_t)_{t\in\mathbb{N}}$ is a stationary ergodic process of functions selected from a finite set $\mathcal{C}$.

C.2 The feasible region for each constraint is bounded: there is a constant $0 < r_{\max} < \infty$ such that for any $c \in \mathcal{C}$ and $\mathbf{r} \in \mathbb{R}_+^N$ satisfying $c(\mathbf{r}) \leq 0$, we have $r_i \leq r_{\max}$ for each $i \in \mathcal{N}$. [1]

C.3 Each function $c \in \mathcal{C}$ is convex and differentiable on an open set containing $[0, r_{\max}]^N$ with $c(\mathbf{0}) \leq 0$ and

$$\min_{\mathbf{r}\in[0,r_{\max}]^N} c(\mathbf{r}) < 0. \tag{2.2}$$

As indicated in Assumption C.1, we model the evolution of the reward constraints is assumed to be stationary ergodic process.Hence, time averages associated with the constraints will converge to their respective statistical averages, and the distribution of the random vector $(C_{t_1+s}, C_{t_2+s}, ..., C_{t_n+s})$ for any choice of indices $t_1, ..., t_n$ does not depend on the shift s, thus the marginal distribution of C_t does not depend on time. We denote the marginal distribution of this process by $(\pi(c))_{c\in\mathcal{C}}$ and let C^π denote a random constraint with this distribution. This model captures a fairly general class of constraints, including, for example, time-varying capacity constraints associated with bandwidth allocation in wireless networks. If condition

[1] We could allow the constant $r_{\max}$ to be user dependent. But, we avoid this for notational simplicity.

C.2 holds, then we can upper bound any feasible allocation under any constraint in $\mathcal{C}$ using $r_{\max}\mathbf{1}_N$ where $\mathbf{1}_N$ is the N length vector with each component equal to one. Condition C.3 ensures that the feasible sets are convex, and the differentiability requirement simplifies the exposition. The remaining requirements in C.3 are useful in studying the optimization problem OPT(T).

Next we introduce the assumptions on the functions $\left(U_i^V\right)_{i\in\mathcal{N}}$ associated with the variability penalties.

Assumptions U.V: (Variability penalty) Let $v_{\max}{:=}r_{\max}^2$.

U.V.1: For each $i \in \mathcal{N}$, U_i^V is well defined and differentiable on an open set containing $[0, v_{\max}]$ satisfying $\min_{v\in[0,v_{\max}]} \left(U_i^V\right)'(v) > 0$, and $\left(U_i^V\right)'(\cdot)$ is Lipschitz continuous.

U.V.2: For each $i \in \mathcal{N}$ and any $z_1, z_2 \in \left[-\sqrt{v_{\max}}, \sqrt{v_{\max}}\right]$ with $z_1 \neq z_2$, and $\alpha \in (0,1)$ with $\bar{\alpha} = 1 - \alpha$, we have

$$U_i^V\left((\alpha z_1 + \bar{\alpha} z_2)^2\right) < \alpha U_i^V\left(z_1^2\right) + \bar{\alpha} U_i^V\left(z_2^2\right). \tag{2.3}$$

The assumptions concerning the Lipschitz continuity of derivatives made in Assumptions U.V.1 and U.E (see below) are made to simplify the exposition, and could be relaxed (see Section 2.5.2). Note that any non-decreasing (not necessarily strictly) convex function satisfies (2.3), but the condition is weaker than a convexity requirement. For instance, using triangle inequality, one can show that $U_i^V(v_i) = \sqrt{v_i + \delta}$ for $\delta > 0$ satisfies all the conditions described above for any $v_{\max}$[2]. This function is not convex but is useful as it transforms variance to (approximately) the standard deviation for small enough $\delta > 0$. We will later see that our algorithm (Section 2.1.1) can be simplified if any of the functions U_i^V are linear. Hence, we define the

[2]Note that we need $\delta > 0$ otherwise $U_i^V(v_i) = \sqrt{v_i}$ violates U.V.1

following subsets of $\mathcal{N}$:

$$\begin{aligned}\mathcal{N}_l &:= \left\{i \in \mathcal{N} : \; U_i^V \text{ is linear}\right\}, \\ \mathcal{N}_n &:= \left\{i \in \mathcal{N} : \; U_i^V \text{ is not linear}\right\}.\end{aligned}$$

Next we discuss assumptions on the functions $\left(U_i^E\right)_{i\in\mathcal{N}}$ used to impose fairness associated with the QoE across users. Recall that our proxy for the QoE for user i is $e_i(t) = m_i(t) - U_i^V\left(v_i(t)\right)$ and, let

$$e_{\min,i} := -U_i^V(v_{\max}) \text{ and } e_{\max,i} := r_{\max} - U_i^V(0).$$

Assumption U.E: (Fairness in QoE)

U.E: For each $i \in \mathcal{N}$, U_i^E is concave and differentiable on an open set containing $[e_{\min,i}, e_{\max,i}]$ with $\left(U_i^E\right)'(e_{\max,i}) > 0$, and $\left(U_i^E\right)'(\cdot)$ is Lipschitz continuous.

Note that concavity and the condition that $\left(U_i^E\right)'(e_{\max,i}) > 0$ ensure that $\left(U_i^E\right)'$ is strictly positive on $[e_{\min,i}, e_{\max,i}]$. For each $i \in \mathcal{N}$, although U_i^E has to be defined over an open set containing $[e_{\min,i}, e_{\max,i}]$, only the definition of the function over $\left[-U_i^V(0), e_{\max,i}\right]$ affects the optimization. This is because we can achieve this value of QoE for each user just by allocating $0\mathbf{1}_N$ in each time slot. Thus, for example, we can choose any function from the following class of strictly concave increasing functions parametrized by $\alpha \in (0,\infty)$ ([34])

$$U_\alpha(e) = \begin{cases}\log(e) & \text{if } \alpha = 1, \\ (1-\alpha)^{-1} e^{1-\alpha} & \text{otherwise,}\end{cases} \tag{2.4}$$

and can satisfy U.E by making minor modifications to the function. For instance, we can use the following modification $U^{E,log}$ of the log function for any (small) $\delta > 0$: $U^{E,log}(e) = \log(e - e_{\min,i} + \delta)$, $e \in [e_{\min,i}, e_{\max,i}]$. The above class of functions are commonly used to enforce fairness specifically to achieve reward allocations that are $\alpha-$fair (see [43]).

Good choices of $\left(U_i^V\right)_{i\in\mathcal{N}}$ and $\left(U_i^E\right)_{i\in\mathcal{N}}$ will depend on the problem setting. A good choice for $\left(U_i^V\right)_{i\in\mathcal{N}}$ should be driven by an understanding of the impact of temporal variability on a user's QoE, which might in turn be based on experimental data. For instance, a choice of $U_i^V(v_i) = \sqrt{v_i + \delta}$ is proposed for video adaptation in [59]. The choice of $\left(U_i^E\right)_{i\in\mathcal{N}}$ is driven by the degree of fairness in the allocation of QoE across users, e.g. max-min, proportional fairness etc. A larger α corresponds to a more fair allocation which eventually becomes max-min fair as α goes to infinity.

Applicability of the model

We close this section by illustrating the wide scope of the framework discussed above by describing examples of scenarios that fit it nicely. They illustrate the freedom provided by the framework for modeling temporal variability in both the available rewards and the sensitivity of the users' reward/utility to their reward allocations, as well as fairness across users' QoE. The presence of time-varying constraints $c_t(\mathbf{r}) \leq 0$ allows us to apply the model to several interesting settings. In particular, we discuss three wireless network settings and show that the framework can handle problems involving time-varying exogenous loads and time-varying utility functions.

Time-varying capacity constraints

We start by discussing the case where the rewards in a time slot is the rate allocated to the users, and users dislike variability in their allocations. Let $\mathcal{P}$ denote a finite (but arbitrarily large) set of positive vectors where each vector corresponds to the peak transmission rates achievable to the set of users in a given time slot. Let

$\mathcal{C} = \left\{ c_{\mathbf{p}} : c_{\mathbf{p}}\left(\mathbf{r}\right) = \sum_{i \in \mathcal{N}} \frac{r_i}{p_i} - 1,\ \mathbf{p} \in \mathcal{P} \right\}$. Here, for any allocation $\mathbf{r}$, r_i/p_i is the fraction of time the wireless system needs to serve user i in time slot t in order to deliver data at the rate of r_i when the user has peak transmission rate p_i. Thus, the constraint $c_{\mathbf{p}}\left(\mathbf{r}\right) \leq 0$ can be seen as a scheduling constraint that corresponds to the requirement that the sum of the fractions of time that different users are served in a time slot should be less than or equal to one.

Time-varying exogenous constraints

We can further introduce time varying exogenous constraints on the wireless system by appropriately defining the set $\mathcal{C}$. For instance, consider a base station in a cellular network that supports users who dislike variability in rate allocation. But, while allocating rates to these users, we may also need to account for the time-varying rate requirements of the voice traffic handled by the base station. We can model this by defining

$$\mathcal{C} = \left\{ c_{\mathbf{p},f} : c_{\mathbf{p},f}\left(\mathbf{r}\right) = \sum_{i \in \mathcal{N}} \frac{r_i}{p_i} - (1-f)\,,\ \mathbf{p} \in \mathcal{P}, f \in \mathcal{T}_{fr} \right\}$$

where $\mathcal{T}_{fr}$ is a finite set of real numbers in $[0, 1)$ where each element in the set corresponds to a fraction of a time slot that is allocated to other traffic.

Time-varying utility functions

Additionally, our framework also allows the introduction of time-varying utility functions as illustrated by the following example of a wireless network supporting video users. Here, we view utility functions as a mapping from allocated resource (e.g., rate) to reward (e.g., video quality). For video users, we consider perceived video quality of a user in a time slot as the reward for that user in that slot. However, for video users, the dependence of perceived video quality [3] on the compression rate is time varying. This is typically due to the possibly changing nature of the content, e.g., from an action to a slower scene. Hence, the utility function that maps the reward (i.e., perceived video quality) derived from the allocated resource (i.e., the rate) is time varying. This setting can be handled as follows. Let $q_{t,i}(\cdot)$ denote the strictly increasing concave function that, in time slot t, maps the rate allocated to user i to user perceived video quality. For each user i, let $\mathcal{Q}_i$ be a finite set of such functions, then a scenario with time varying peak rates and utilities can be modeled by set of convex constraints:

$$\mathcal{C} = \left\{ c_{\mathbf{p},\mathbf{q}} : c_{\mathbf{p},\mathbf{q}}(\mathbf{r}) = \sum_{i \in \mathcal{N}} \frac{q_i^{-1}(r_i)}{p_i} - 1,\ \mathbf{p} \in \mathcal{P},\ q_i \in \mathcal{Q}_i \quad \forall\, i \in \mathcal{N} \right\}.$$

2.3 Optimal variance-sensitive offline policy

In this section, we study OPT(T), the offline formulation for optimal reward allocation introduced in Section 2.1. In the offline setting, we assume that $(c)_{1:T}$, the realization of the constraints process $(C)_{1:T}$, is known. We denote the objective function of OPT(T) by ϕ_T, i.e.,

$$\phi_T(\mathbf{r}) := \sum_{i \in \mathcal{N}} U_i^E\left(e_i^T(r_i)\right), \tag{2.5}$$

[3]in a short duration time slot roughly a second long which corresponds to a collection of 20-30 frames

where $e_i^T(\cdot)$ is as in Definition 2.1. Hence the optimization problem OPT(T) can be rewritten as:

$$\max_{(\mathbf{r})_{1:T}} \quad \phi_T(\mathbf{r}) \tag{2.6}$$

$$\text{subject to} \quad c_t(\mathbf{r}(t)) \leq 0 \;\; \forall\, t \in \{1, ..., T\}, \tag{2.7}$$

$$r_i(t) \geq 0 \quad \forall\, t \in \{1, ..., T\}, \forall\, i \in \mathcal{N}. \tag{2.8}$$

The next result asserts that OPT(T) is a convex optimization problem satisfying Slater's condition (Section 5.2.3, [10]) and that it has a unique solution.

Lemma 2.1. *OPT(T) is a convex optimization problem satisfying Slater's condition with a unique solution.*

Proof. By Assumptions U.E and U.V, the convexity of the objective of OPT(T) is easy to establish once we prove the convexity of the function $U_i^V(\mathrm{Var}^T(\cdot))$ for each $i \in \mathcal{N}$. Using (2.3) and the definition of $\mathrm{Var}^T(\cdot)$, we can show that $U_i^V(\mathrm{Var}^T(\cdot))$ is a convex function for each $i \in \mathcal{N}$. The details are given next. For any two quality vectors $\left(\mathbf{r^1}\right)_{1:T}$ and $\left(\mathbf{r^2}\right)_{1:T}$, any $i \in \mathcal{N}$, $\alpha \in (0,1)$ and $\bar{\alpha} = 1 - \alpha$, we have that

$$\begin{aligned}
&\mathrm{Var}^T\left(\alpha r_i^1 + \bar{\alpha} r_i^2\right) \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\alpha\left(r_i^1(t) - m^T\left(r_i^1\right)\right) + \bar{\alpha}\left(r_i^2(t) - m^T\left(r_i^2\right)\right)\right)^2 \\
&\leq \left(\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\alpha\left(r_i^1(t) - m^T\left(r_i^1\right)\right)\right)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\bar{\alpha}\left(r_i^2(t) - m^T\left(r_i^2\right)\right)\right)^2}\right)^2 \\
&= \left(\alpha\sqrt{\mathrm{Var}^T\left(r_i^1\right)} + \bar{\alpha}\sqrt{\mathrm{Var}^T\left(r_i^2\right)}\right)^2
\end{aligned} \tag{2.9}$$

where the above inequality follows from triangle inequality for the Euclidean norm.

Using this, (2.3) and the monotonicity of U_i^V, we have

$$U_i^V\left(\mathrm{Var}^T\left(\alpha r_i^1+\bar{\alpha} r_i^2\right)\right) \leq \alpha U_i^V\left(\mathrm{Var}^T\left(r_i^1\right)\right)+\bar{\alpha} U_i^V\left(\mathrm{Var}^T\left(r_i^2\right)\right). \tag{2.10}$$

So, $U_i^V\left(\mathrm{Var}^T\left(\cdot\right)\right)$ is a convex function. Thus, by the concavity of $U_i^E\left(\cdot\right)$ and $-U_i^V(\mathrm{Var}^T\left(\cdot\right))$, we can conclude that OPT(T) is a convex optimization problem. Also, from (2.9) and (2.3) (since we have strict inequality), we can conclude that we have equality in (2.10) only if

$$\mathrm{Var}^T\left(r_i^1\right)=\mathrm{Var}^T\left(r_i^2\right), \tag{2.11}$$

or equivalently

$$r_i^1(t)=r_i^2(t)+m^T\left(r_i^1\right)-m^T\left(r_i^2\right) \quad \forall t \in\{1,...,T\}. \tag{2.12}$$

Further, Slater's condition is satisfied and it follows from (2.2) in Assumption C.3.

Now, for any $i \in \mathcal{N}$, U_i^E and $-U_i^V(\mathrm{Var}^T\left(\cdot\right))$ are not necessarily strictly concave. But, we can still show that OPT(T) has a unique solution. Let $\left(\mathbf{r^1}\right)_{1:T}$ and $\left(\mathbf{r^2}\right)_{1:T}$ be two optimal solutions to OPT(T). Then, from the concavity of the objective, $\left(\alpha\left(r_i^1\right)_{1:T}+\bar{\alpha}\left(r_i^2\right)_{1:T}\right)$ is also an optimal solution for any $\alpha \in(0,1)$ and $\bar{\alpha}=1-\alpha$. Due to convexity of $U_i^E\left(\cdot\right)$ and $U_i^V\left(\mathrm{Var}^T\left(\cdot\right)\right)$, this is only possible if for each $i \in \mathcal{N}$ and $1 \leq t \leq T$,

$$U_i^V\left(\mathrm{Var}^T\left(\alpha r_i^1+\bar{\alpha} r_i^2\right)\right)=\alpha U_i^V\left(\mathrm{Var}^T\left(r_i^1\right)\right)+\bar{\alpha} U_i^V\left(\mathrm{Var}^T\left(r_i^2\right)\right).$$

Hence (2.12) and (2.11) hold. Due to optimality of $\left(\mathbf{r^1}\right)_{1:T}$ and $\left(\mathbf{r^2}\right)_{1:T}$, we have that

$$
\begin{aligned}
&\sum_{i\in\mathcal{N}} U_i^E \left(\frac{1}{T}\sum_{t=1}^{T} r_i^2(t) - U_i^V\left(\mathrm{Var}^T\left(r_i^2\right)\right)\right) \\
&\quad = \sum_{i\in\mathcal{N}} U_i^E \left(\frac{1}{T}\sum_{t=1}^{T} r_i^1(t) - U_i^V\left(\mathrm{Var}^T\left(r_i^2\right)\right)\right) \\
&\quad = \sum_{i\in\mathcal{N}} U_i^E \left(\frac{1}{T}\sum_{t=1}^{T} r_i^2(t) + m^T\left(r_i^1\right) - m^T\left(r_i^2\right) - U_i^V\left(\mathrm{Var}^T\left(r_i^2\right)\right)\right),
\end{aligned}
$$

where the first equality follows from (2.11) and the second one follows from (2.12). Since U_i^E is a strictly increasing function for each $i \in \mathcal{N}$, the above equation implies that $m^T\left(r_i^1\right) = m^T\left(r_i^2\right)$ and thus (using (2.12)) $\mathbf{r^1}(t) = \mathbf{r^2}(t)$ for each t such that $1 \le t \le T$. From the above discussion, we can conclude that OPT(T) has a unique solution. □

We let $\left(\mathbf{r}^T\right)_{1:T}$ denote the optimal solution to OPT(T). Since OPT(T) is a convex optimization problem satisfying Slater's condition (Lemma 2.1), the Karush-Kuhn-Tucker (KKT) conditions (see Section 5.5.3 in [10]) given next hold.

KKT-OPT(T):

There exist nonnegative constants $\left(\mu^T\right)_{1:T}$ and $\left(\boldsymbol{\gamma}^T\right)_{1:T}$ such that for all $i \in \mathcal{N}$ and $t \in \{1, ..., T\}$, we have

$$
\begin{aligned}
\left(U_i^E\right)'\left(e_i^T\left(r_i^T\right)\right)&\left(\frac{1}{T} - \frac{2\left(U_i^V\right)'\left(\mathrm{Var}^T\left(r_i^T\right)\right)}{T}\left(r_i^T(t) - m^T\left(r_i^T\right)\right)\right) \\
&- \frac{\mu^T(t)}{T} c'_{t,i}(\mathbf{r}^T(t)) + \frac{\gamma_i^T(t)}{T} = 0, \qquad (2.13) \\
&\mu^T(t) c_t(\mathbf{r}^T(t)) = 0, \qquad (2.14) \\
&\gamma_i^T(t) r_i^T(t) = 0, \qquad (2.15)
\end{aligned}
$$

Here $c'_{t,i}$ denotes $\frac{\partial c_t}{\partial r_i}$, and we have used the fact that for any $t \in \{1, ..., T\}$

$$\frac{\partial}{\partial r(t)} \left(T\mathrm{Var}^T(r) \right) = 2 \left(r(t) - m^T(r) \right).$$

From (2.13), we see that the optimal reward allocation $\mathbf{r}^T(t)$ on time slot t depends on the entire allocation $\left(\mathbf{r^T}\right)_{1:T}$ through the following three quantities: (i) the time average rewards $\mathbf{m}^T$; (ii) $\left(\left(U_i^E\right)'\right)_{i\in\mathcal{N}}$ evaluated at the quality of experience of the respective users; and (iii), $\left(\left(U_i^V\right)'\right)_{i\in\mathcal{N}}$ evaluated at the variance seen by the respective users. So, if the time averages associated with the *optimal solution* were somehow known, the optimal allocation for each time slot t could be determined by solving an optimization problem (derived from the KKT conditions) that only requires these time averages, and knowledge of c_t (associated with current time slot) rather than $(c)_{1:T}$. We exploit this key idea in formulating our online algorithm in the next section.

2.4 Adaptive variance-aware reward allocation

In this section, we present Adaptive Variance-aware Reward allocation (AVR) algorithm and establish its asymptotic optimality.

We let

$$\mathcal{H} \quad := \quad [0, r_{\max}]^N \times [0, v_{\max}]^N, \tag{2.16}$$

where $\times$ denotes Cartesian product for sets. Let $(\mathbf{m}, \boldsymbol{v}) \in \mathcal{H}$ and $e_i = m_i - U_i^V(v_i)$ for each $i \in \mathcal{N}$, and consider the optimization problem OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$ given below:

$$\begin{aligned} \max_{\mathbf{r}} \sum_{i\in\mathcal{N}} & \left(U_i^E\right)'(e_i)\left(r_i - \left(U_i^V\right)'(v_i)\left(r_i - m_i\right)^2\right) \\ \text{subject to} \quad & c(\mathbf{r}) \le 0, && (2.17) \\ & r_i \ge 0 \quad \forall\, i \in \mathcal{N}. && (2.18) \end{aligned}$$

The reward allocations for AVR are obtained by solving OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$, where $\mathbf{m}$, $\boldsymbol{v}$ and $\mathbf{e}$ correspond to current estimates of the mean, variance and QoE respectively. We let $\mathbf{r}^*((\mathbf{m}, \boldsymbol{v}), c)$ denote the optimal solution to OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$.

Next, we describe our algorithm in detail.

Algorithm 2.1. Adaptive Variance-aware Reward allocation (AVR)

AVR.0: Initialization: let $(\mathbf{m}(1), \boldsymbol{v}(1)) \in \mathcal{H}$.

In each time slot $t \in \mathbb{N}$, carry out the following steps:

AVR.1: The reward allocation in time slot t is $\mathbf{r}^*((\mathbf{m}(t), \boldsymbol{v}(t)), c_t)$, i.e., the optimal solution to OPTAVR$((\mathbf{m}(t), \boldsymbol{v}(t)), c_t)$, and will be denoted by $\mathbf{r}^*(t)$ (when the dependence on the variables is clear from context).

AVR.2: In time slot t, update m_i as follows: for all $i \in \mathcal{N}$,

$$m_i(t+1) = \left[m_i(t) + \frac{1}{t}\left(r_i^*(t) - m_i(t)\right)\right]_0^{r_{\max}}, \tag{2.19}$$

and update v_i as follows: for all $i \in \mathcal{N}$,

$$v_i(t+1) = \left[v_i(t) + \frac{\left(r_i^*(t) - m_i(t)\right)^2 - v_i(t)}{t}\right]_0^{v_{\max}}. \tag{2.20}$$

Here, $[x]_a^b = \min\left(\max\left(x, a\right), b\right)$.

Thus, AVR greedily allocates rewards in slot t based on the objective of OPTAVR$((\mathbf{m}(t), \boldsymbol{v}(t)), c_t)$. Thus, the computational requirements per slot involve solving a convex program in N variables (that has a simple quadratic function as its objective function), and updating at most $2N$ variables. We see that the update equations (2.19)-(2.20) roughly ensure that the parameters $\mathbf{m}(t)$ and $\boldsymbol{v}(t)$ keep track of mean reward and variance in reward allocations under AVR. The updates in AVR fall in the class of decreasing step size stochastic approximation algorithms (see [29] for reference) due to the use of $1/t$ in (2.19)-(2.20). We could replace $1/t$ with a small positive constant ϵ and obtain a constant step size stochastic approximation algorithm which is usually better suited for non-stationary settings. Also, note that we do not have to keep track of variance estimates for users i with linear U_i^V since OPTAVR is insensitive to their values (i.e., $\left(U_i^V\right)'(.)$ is a constant), and thus the evolutions of $\mathbf{m}(t)$ and $(v_i(t))_{i\in\mathcal{N}_n}$ do not depend on them. We let $\boldsymbol{\theta}(t) = (\mathbf{m}(t), \boldsymbol{v}(t))$ for each t. The truncation $[.]_a^b$ in the update equations (2.19)-(2.20) ensure that $\boldsymbol{\theta}(t)$ stays in the set $\mathcal{H}$.

For any $((\mathbf{m}, \boldsymbol{v}), c) \in \mathcal{H} \times \mathcal{C}$, we have $\left(U_i^E\right)'\left(m_i - U_i^V(v_i)\right)\left(U_i^V\right)'(v_i) > 0$ for each $i \in \mathcal{N}$ (see Assumptions U.E and U.V). Hence, OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$ is a convex optimization problem with a unique solution. Further, using (2.2) in Assumption C.3, we can show that it satisfies Slater's condition. Hence, the optimal solution $\mathbf{r}^*$ for OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$ satisfies KKT conditions given below.

KKT-OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$:

There exist nonnegative constants μ^* and $(\gamma_i^*)_{i\in\mathcal{N}}$ such that for all $i \in \mathcal{N}$

$$\left(U_i^E\right)'\left(m_i - U_i^V(v_i)\right)\left(1 - 2\left(U_i^V\right)'(v_i)(r_i^* - m_i)\right) + \gamma_i^* - \mu^* c_i'(\mathbf{r}^*) = 0, \quad (2.21)$$

$$\mu^* c(\mathbf{r}^*) = 0, \quad (2.22)$$

$$\gamma_i^* r_i^* = 0. \quad (2.23)$$

In the next lemma, we establish continuity properties of $\mathbf{r}^*((\mathbf{m}, \boldsymbol{v}), c)$ when viewed as a function of $(\mathbf{m}, \boldsymbol{v})$. In particular, the Lipschitz assumption on the derivatives of $\left(U_i^V\right)_{i\in\mathcal{N}}$ and $\left(U_i^E\right)_{i\in\mathcal{N}}$ help us conclude that the optimizer of OPTAVR($\boldsymbol{\theta}, c$) is Lipschitz continuous in $\boldsymbol{\theta}$.

Lemma 2.2. *For any $c \in \mathcal{C}$, and $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{v}) \in \mathcal{H}$*

(a) $\mathbf{r}^*(\boldsymbol{\theta}, c)$ is a Lipschitz continuous function of $\boldsymbol{\theta}$.

(b) $E[\mathbf{r}^*(\boldsymbol{\theta}, C^\pi)]$ is a Lipschitz continuous function of $\boldsymbol{\theta}$.

Proof. For $\boldsymbol{\theta} = (\mathbf{m}, \mathbf{v})$, let

$$\Phi_{\boldsymbol{\theta}}(\mathbf{r}) := \sum_{i\in\mathcal{N}} \left(U_i^E\right)'(e_i)\left(r_i - \left(U_i^V\right)'(v_i)(r_i - m_i)^2\right) \tag{2.24}$$

for $\mathbf{r} \in \mathbb{R}^N$ where $e_i = m_i - U_i^V(v_i)$ for each $i \in \mathcal{N}$. Next, for any $\boldsymbol{\theta}^a, \boldsymbol{\theta}^b \in \mathcal{H}$ and $\mathbf{r} \in [-2r_{\max}, 2r_{\max}]^N$ (any optimal solution to OPTAVR, i.e., minimizer of $\Phi_{\boldsymbol{\theta}}(\mathbf{r})$ subject to constraints is an interior point of this set), let

$$\Delta\Phi\left(\mathbf{r}, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b\right) = \Phi_{\boldsymbol{\theta}^b}(\mathbf{r}) - \Phi_{\boldsymbol{\theta}^a}(\mathbf{r}).$$

We prove part (a) (i.e., the Lipschitz continuity with respect to $\boldsymbol{\theta}$ of the optimizer $\mathbf{r}^*(\boldsymbol{\theta}, c)$ of $\Phi_{\boldsymbol{\theta}}(\mathbf{r})$ subject to constraint c) using Proposition 4.32 in [8]. The first condition in the Proposition requires that $\Delta\Phi\left(\cdot, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b\right)$ be Lipschitz continuous. To show this, note that for any $\mathbf{r}^c, \mathbf{r}^d \in [-2r_{\max}, 2r_{\max}]^N$

$$
\begin{aligned}
&\Delta\Phi\left(\mathbf{r}^c, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b\right) - \Delta\Phi\left(\mathbf{r}^d, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b\right) \\
&= \sum_{i\in\mathcal{N}} \left(\left(U_i^E\right)'\left(e_i^a\right) - \left(U_i^E\right)'\left(e_i^b\right)\right)\left(r_i^c - r_i^d\right) \\
&\quad + \sum_{i\in\mathcal{N}} \left(U_i^E\right)'\left(e_i^a\right)\left(U_i^V\right)'\left(v_i^a\right)\left(r_i^d - r_i^c\right)\left(r_i^d + r_i^c - 2m_i^a\right) \\
&\quad - \sum_{i\in\mathcal{N}} \left(U_i^E\right)'\left(e_i^b\right)\left(U_i^V\right)'\left(v_i^b\right)\left(r_i^d - r_i^c\right)\left(r_i^d + r_i^c - 2m_i^b\right).
\end{aligned}
$$

Using the above expression, Lipschitz continuity and boundedness of $\left(U_i^{V\prime}\right)_{i\in\mathcal{N}}$ and $\left(U_i^{E\prime}\right)_{i\in\mathcal{N}}$ (see Assumptions U.V.1 and U.E), and boundedness of $\mathbf{r}^a$ and $\mathbf{r}^b$, we can conclude that there exists some positive finite constant η such that

$$
\Delta\Phi\left(\mathbf{r}^c, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b\right) \le \eta d\left(\boldsymbol{\theta}^a, \boldsymbol{\theta}^b\right) d\left(\mathbf{r}^a, \mathbf{r}^b\right).
$$

Next, we establish the second condition given in the proposition referred to as second order growth condition. For this we use Theorem 6.1 (vi) from [6], and consider the functions L and ψ discussed in the exposition of the theorem. We have

$$
L\left(\mathbf{r}, \boldsymbol{\theta}, \mu, \boldsymbol{\gamma}, c\right) = \Phi_{\boldsymbol{\theta}}\left(\mathbf{r}^*\right) - \Phi_{\boldsymbol{\theta}}\left(\mathbf{r}\right) + \mu c\left(\mathbf{r}\right) - \sum_{i\in\mathcal{N}} \gamma_i r_i,
$$

and for $d \in \mathbb{R}^{\mathcal{N}}$, we have

$$
\psi_{\mathbf{r}^*(\boldsymbol{\theta}^a, c)}\left(\mathbf{d}\right) = \mathbf{d}^{tr}\nabla_{\mathbf{r}}^2 L\left(\mathbf{r}^*\left(\boldsymbol{\theta}^a, c\right), \boldsymbol{\theta}^a, \mu^m(c), \boldsymbol{\gamma}^m(c), c\right)\mathbf{d}
$$

where $\mu^m(c)$ and $(\gamma_i^m(c) : i \in \mathcal{N})$ are Lagrange multipliers associated with the optimal solution to OPTAVR$(\boldsymbol{\theta}^a, c)$. Then, using convexity of c we have

$$
\psi_{\mathbf{r}^*(\boldsymbol{\theta}^a, c)}\left(\mathbf{d}\right) \ge \sum_{i\in\mathcal{N}} 2\left(U_i^E\right)'\left(e_i^a\right)\left(U_i^V\right)'\left(v_i^a\right) d_i^2.
$$

Since $\left(U_i^{V\prime}\right)_{i\in\mathcal{N}}$ and $\left(U_i^{E\prime}\right)_{i\in\mathcal{N}}$ are strictly positive (see Assumptions U.V.1 and U.E), we can conclude that there exists some positive finite constant η_1 such that $\psi_{\mathbf{r}^*(\boldsymbol{\theta}^a,c)}(\mathbf{d}) \geq \eta_1 \|\mathbf{d}\|^2$. Now, using Theorem 6.1 (vi) from [6], we can conclude that second order growth condition is satisfied.

Thus, we have verified the conditions given in Proposition 4.32 in [8], and thus (a) holds. Then, (b) follows from (a) since $\mathcal{C}$ is finite and

$$E\left[\mathbf{r}^*\left(\boldsymbol{\theta}, C^\pi\right)\right] = \sum_{c\in\mathcal{C}} \pi(c)\mathbf{r}^*\left(\boldsymbol{\theta}, c\right).$$

□

2.4.1 Proof of Theorem 2.3

Proof. By KKT-OPTSTAT $(\boldsymbol{\rho}_c^\pi : c \in \mathcal{C})$, $(\mu^\pi(c) : c \in \mathcal{C})$ and $\left(\left(\gamma_i^\pi(c)\right)_{i\in\mathcal{N}} : c \in \mathcal{C}\right)$ satisfy (2.30)-(2.32). To show that $\mathbf{r}^*\left(\left(\mathbf{m}^\pi, \boldsymbol{v}^\pi\right), c\right) = \boldsymbol{\rho}_c^\pi$, we verify that $\boldsymbol{\rho}_c^\pi$ satisfies KKT-OPTAVR$(\left(\mathbf{m}^\pi, \boldsymbol{v}^\pi\right), c)$. To that end, we can verify that $\boldsymbol{\rho}_c^\pi$ along with $\mu^* = \frac{\mu^\pi(c)}{\pi(c)}$ and $\left(\gamma_i^* = \frac{\gamma_i^\pi(c)}{\pi(c)} : i \in \mathcal{N}\right)$ satisfy (2.21)-(2.23) by using (2.30)-(2.32). This proves part (a).

To prove part (b), first note that $(\mathbf{m}^\pi, \boldsymbol{v}^\pi) \in \mathcal{H}^*$ and this follows from (a) and the definitions (see (2.33)-(2.34)) of $\mathbf{m}^\pi$ and $\boldsymbol{v}^\pi$. Next, note that for any $(\mathbf{m}, \boldsymbol{v}) \in \mathcal{H}^*$ and each $c \in \mathcal{C}$, $\mathbf{r}^*(\mathbf{m}, \boldsymbol{v}, c)$ is an optimal solution to OPTAVR and thus, there exist nonnegative constants $\mu^*(c)$ and $(\gamma_i^*(c) : i \in \mathcal{N})$ such that for all $i \in \mathcal{N}$, and satisfies KKT-OPTAVR given in (2.21)-(2.23). Also, since $(\mathbf{m}, \boldsymbol{v}) \in \mathcal{H}^*$, it satisfies

(2.36)-(2.37). Combining these observations, we have that for all $c \in \mathcal{C}$

$$\begin{aligned}
\left(U_i^E\right)' \left(E\left[\mathbf{r}^*\left(\boldsymbol{\theta}, C^\pi\right)\right] - U_i^V\left(\mathrm{Var}^\pi\left(\mathbf{r}^*\left(\boldsymbol{\theta}, C^\pi\right)\right)\right)\right) & \\
\Big(r_i^*\left(\boldsymbol{\theta}, c\right) - 2\left(U_i^V\right)'\left(\mathrm{Var}^\pi\left(\mathbf{r}^*\left(\boldsymbol{\theta}, C^\pi\right)\right)\right)\left(r_i^*\left(\boldsymbol{\theta}, c\right)\right. & \\
\left.-E\left[\mathbf{r}^*\left(\boldsymbol{\theta}, C^\pi\right)\right]\right)\Big) + \gamma_i^* - \mu^*\left(c\right)c_i'(\mathbf{r}^*\left(\boldsymbol{\theta}, c\right)) &= 0, \\
\mu^*\left(c\right)c(\mathbf{r}^*\left(\boldsymbol{\theta}, c\right)) &= 0, \\
\gamma_i^* r_i^*\left(\boldsymbol{\theta}, c\right) &= 0.
\end{aligned}$$

where $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{v})$, and $e_i = m_i - U_i^V\left(v_i\right)$ for each $i \in \mathcal{N}$. Now for each $c \in \mathcal{C}$, multiply the above equations with $\pi(c)$ and one obtains KKT-OPTSTAT ((2.30)-(2.32)) with $\left(\pi\left(c\right)\mu^*\left(c\right) : c \in \mathcal{C}\right)$ and $\left(\left(\pi\left(c\right)\gamma_i^*\left(c\right)\right)_{i\in\mathcal{N}} : c \in \mathcal{C}\right)$ as associated Lagrange multipliers. From Lemma 2.3, OPTSTAT satisfies Slater's condition and hence satisfying KKT conditions is sufficient for optimality for OPTSTAT. Thus, we have that $\left(\mathbf{r}^*\left(\mathbf{m}, \boldsymbol{v}, c\right)\right)_{c\in\mathcal{C}}$ is an optimal solution to OPTSTAT. This observation along with uniqueness of solution to OPTSTAT and (2.36)-(2.37), imply part (b), i.e., $\mathcal{H}^* = \{(\mathbf{m}^\pi, \boldsymbol{v}^\pi)\}$. □

The next theorem states our key convergence result for the mean, variance and QoE of the reward allocations under AVR. This result is proven in Section 2.5. For brevity, we let $\mathbf{r}^*(t)$ denote $\mathbf{r}^*\left(\left(\mathbf{m}(t), \boldsymbol{v}(t)\right), c_t\right)$.

Theorem 2.1. *The evolution of the users' estimated parameters $\mathbf{m}(t)$ and $\mathbf{v}(t)$, and the sequence of reward allocations $(r_i^*)_{1:T}$ to each user i under AVR satisfy the following property: for almost all sample paths, and for each $i \in \mathcal{N}$,*

$$\begin{aligned}
&(a) \quad \lim_{T\to\infty} m^T\left(r_i^*\right) = \lim_{t\to\infty} m_i(t), \\
&(b) \quad \lim_{T\to\infty} \mathrm{Var}^T\left(r_i^*\right) = \lim_{t\to\infty} v_i(t), \\
&(c) \quad \lim_{T\to\infty} e_i^T\left(r_i^*\right) = \lim_{t\to\infty} \left(m_i(t) - U_i^V\left(v_i(t)\right)\right).
\end{aligned}$$

The next result establishes the asymptotic optimality of AVR, i.e., if we consider long periods of time T, the difference in performance (i.e., ϕ_T defined in (2.5)) of the online algorithm AVR and the optimal offline policy OPT(T) becomes negligible. Thus, the sum utility of the QoEs (which depends on long term time averages) is optimized.

Theorem 2.2. *The sequence of reward allocations $(\mathbf{r}^*)_{1:T}$ under AVR is feasible, i.e., it satisfies* (2.7) *and* (2.8)*, and for almost all sample paths they are asymptotically optimal, i.e.,*

$$\lim_{T\to\infty}\left(\phi_T\left(\mathbf{r}^*\right)-\phi_T\left(\mathbf{r^T}\right)\right)=0.$$

Proof. Since the allocation $(\mathbf{r}^*)_{1:T}$ associated with AVR satisfies (2.17) and (2.18) at each time slot, it also satisfies (2.7) and (2.8).

To show asymptotic optimality, consider any realization of $(c)_{1:T}$. Let $(\mu^*)_{1:T}$ and $(\boldsymbol{\gamma}^*)_{1:T}$ be the sequences of nonnegative real numbers satisfying (2.21), (2.22) and (2.23) for this realization. From the nonnegativity of these numbers, and feasibility of $\left(\mathbf{r}^T\right)_{1:T}$, we have

$$\phi_T\left(\mathbf{r}^T\right) \quad\leq\quad \psi_T\left(\mathbf{r}^T\right), \tag{2.25}$$

where

$$\psi_T\left(\mathbf{r}^T\right) \quad=\quad \sum_{i\in\mathcal{N}}U_i^E\left(e_i^T\left(r_i^T\right)\right)-\sum_{t=1}^{T}\frac{\mu^*(t)}{T}c_t(\mathbf{r}^T(t))+\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\frac{\gamma_i^*(t)}{T}r_i^T(t).$$

Indeed, the function ψ_T is the Lagrangian associated with OPT(T) but evaluated at the optimal Lagrange multipliers associated with the optimization problems (OP-

TAVR) involved in AVR, and hence the inequality. Since ψ_T is a differentiable concave function, we have (see [10])

$$\psi_T\left(\mathbf{r}^T\right) \leq \psi_T\left(\mathbf{r}^*\right) + \left\langle \nabla\psi_T\left(\mathbf{r}^*\right), \left(\left(\mathbf{r}^T\right)_{1:T} - \left(\mathbf{r}^*\right)_{1:T}\right)\right\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the dot product. Hence, we have

$$\begin{aligned}
\psi_T\left(\mathbf{r}^T\right) \quad \leq \quad & \sum_{i\in\mathcal{N}} U_i^E\left(e_i^T\left(r_i^*\right)\right) - \sum_{t=1}^{T}\frac{\mu^*(t)}{T}c_t(\mathbf{r}^*(t)) + \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\frac{\gamma_i^*(t)}{T}r_i^*(t) \\
& + \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\left(r_i^T(t) - r_i^*(t)\right) \\
& \qquad \left(-\frac{\mu^*(t)}{T}c'_{t,i}(\mathbf{r}^*(t)) + \frac{\gamma_i^*(t)}{T} + \left(U_i^E\right)'\left(e_i^T\left(r_i^*\right)\right)\right. \\
& \qquad \left.\left(\frac{1}{T} - \frac{2\left(U_i^V\right)'\left(\mathrm{Var}^T\left(r_i^*\right)\right)}{T}\left(r_i^*(t) - m^T\left(r_i^*\right)\right)\right)\right).
\end{aligned}$$

Using (2.25), and the fact that $(\mu^*)_{1:T}$ and $(\boldsymbol{\gamma}^*)_{1:T}$ satisfy (2.21), (2.22) and (2.23), we have

$$\begin{aligned}
\phi_T\left(\mathbf{r}^T\right) \leq & \sum_{i\in\mathcal{N}} U_i^E\left(e_i^T\left(r_i^*\right)\right) + \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\frac{r_i^T(t) - r_i^*(t)}{T} \\
& \Bigg(\left(U_i^E\right)'\left(e_i^T\left(r_i^*\right)\right)\left(1 - 2\left(U_i^V\right)'\left(\mathrm{Var}^T\left(r_i^*\right)\right)\left(r_i^*(t) - m^T\left(r_i^*\right)\right)\right) \\
& - \left(U_i^E\right)'\left(e_i(t-1)\right)\left(1 - 2\left(U_i^V\right)'\left(v_i(t-1)\right)\left(r_i^*(t) - m_i(t-1)\right)\right)\Bigg).
\end{aligned} \tag{2.26}$$

From Theorem 2.1 (a)-(c), and the continuity and boundedness of the functions involved, we can conclude that the expression appearing in the last four lines of the above inequality can be made as small as desired by choosing large enough T and then choosing a large enough t. Also, $\left|r_i^T(t) - r_i^*(t)\right| \leq r_{\max}$ for each $i \in \mathcal{N}$. Hence, taking limits in (2.26),

$$\lim_{T\to\infty}\left(\phi_T\left(\mathbf{r}^*\right)-\phi_T\left(\mathbf{r}^T\right)\right)\geq 0. \tag{2.27}$$

holds for almost all sample paths. From the optimality of $\left(\mathbf{r}^T\right)_{1:T}$,

$$\phi_T\left(\mathbf{r}^T\right)\geq \phi_T\left(\mathbf{r}^*\right). \tag{2.28}$$

The result follows from the inequalities (2.27) and (2.28). □

2.5 Convergence analysis

This section is devoted to the proof of the previously stated Theorem 2.1 capturing the convergence of reward allocations under AVR. Our approach relies on viewing (2.19)-(2.20) in AVR as a stochastic approximation update equation (see, e.g., [29] for reference), and relating the convergence of reward allocations under the discrete time algorithm AVR to that of an auxiliary (continuous time) ODE (given in (2.38)) which evolves according to time averaged dynamics of AVR. In fact, we will show that the ODE converges to a point determined by the optimal solution to an auxiliary optimization problem OPTSTAT closely related to OPT(T) which is discussed in the next subsection. In Subsection 2.5.2, we study the convergence of the auxiliary ODE and in Subsection 2.5.3, we establish convergence of $(\boldsymbol{\theta}(t))_{t\in\mathbb{N}}$ generated by AVR to complete the proof of Theorem 2.1.

2.5.1 A stationary version of OPT: OPTSTAT

The formulation OPT(T) involves time averages of various quantities associated with users' rewards. By contrast, the formulation of OPTSTAT is based on expected values of the corresponding quantities under the stationary distribution of $(C_t)_{t\in\mathbb{N}}$.

Recall that (under Assumption C.1) $(C_t)_{t\in\mathbb{N}}$ is a stationary ergodic process with marginal distribution $(\pi(c))_{c\in\mathcal{C}}$, i.e., for $c \in \mathcal{C}$, $\pi(c)$ is the probability of the event $C_t = c$. Since $\mathcal{C}$ is finite, we assume that $\pi(c) > 0$ for each $c \in \mathcal{C}$ without any loss of generality.

Definition 2.2. A reward allocation policy is said to be *stationary* if the associated reward allocation in any time slot t depends only on current constraint c_t.

Thus, we can represent any stationary reward allocation policy as a $|\mathcal{C}|$ length vector (of vectors) $(\boldsymbol{\rho}_c)_{c\in\mathcal{C}}$ where $\boldsymbol{\rho}_c = (\rho_{c,i})_{i\in\mathcal{N}} \in \mathbb{R}_+^N$ denotes the allocation of rewards to users under constraint $c \in \mathcal{C}$.

Definition 2.3. We say that a stationary reward allocation policy $(\boldsymbol{\rho}_c)_{c\in\mathcal{C}}$ is *feasible* if for each $c \in \mathcal{C}$, we have that $c(\boldsymbol{\rho}_c) \leq 0$ and for each $i \in \mathcal{N}$, we have $\rho_{c,i} \geq 0$. Also, let $\mathcal{R}_{\mathcal{C}} \subset \mathbb{R}^{N|\mathcal{C}|}$ denote the set of feasible stationary reward allocation policies, i.e.,

$$\mathcal{R}_{\mathcal{C}} {:=} \Pi_{c\in\mathcal{C}} \left\{ \boldsymbol{\rho}_c \in \mathbb{R}^N : c(\boldsymbol{\rho}_c) \leq 0,\ \rho_{c,i} \geq 0 \quad \forall\, i \in \mathcal{N} \right\}. \tag{2.29}$$

Now, let

$$\phi_\pi\left((\boldsymbol{\rho}_c)_{c\in\mathcal{C}}\right) = \sum_{i\in\mathcal{N}} U_i^E\left(E\left[\rho_{C^\pi,i}\right] - U_i^V\left(\mathrm{Var}\left(\rho_{C^\pi,i}\right)\right)\right)$$

where $\rho_{C^\pi,i}$ is a random variable taking value $\rho_{c,i}$ with probability $\pi(c)$ for each $c \in \mathcal{C}$, i.e., a random variable whose distribution is that of user i's reward allocation under stationary reward allocation policy $(\boldsymbol{\rho}_c)_{c\in\mathcal{C}}$. Hence,

$$\begin{aligned} E\left[\rho_{C^\pi,i}\right] &= \sum_{c\in\mathcal{C}} \pi(c)\rho_{c,i}, \\ \mathrm{Var}\left(\rho_{C^\pi,i}\right) &= \sum_{c\in\mathcal{C}} \pi(c)\left(\rho_{c,i} - E\left[\rho_{C^\pi,i}\right]\right)^2. \end{aligned}$$

We define the 'stationary' optimization problem OPTSTAT as follows:

$$\max_{(\boldsymbol{\rho}_c)_{c\in\mathcal{C}}\in\mathcal{R}_{\mathcal{C}}} \qquad \phi_\pi\left((\boldsymbol{\rho}_c)_{c\in\mathcal{C}}\right).$$

The next lemma gives a few useful properties of OPTSTAT.

Lemma 2.3. *OPTSTAT is a convex optimization problem satisfying Slater's condition and has a unique solution.*

Proof. The proof is similar to that of Lemma 2.1, and is easy to establish once the convexity of the function $\mathrm{Var}\left(\cdot\right)$ is shown. □

Using Lemma 2.3, we can conclude that the KKT conditions given below are necessary and sufficient for optimality of OPTSTAT. Let $(\boldsymbol{\rho}_c^\pi)_{c\in\mathcal{C}}$ denote the optimal solution.

KKT-OPTSTAT:

There exist constants $(\mu^\pi\left(c\right))_{c\in\mathcal{C}}$ and $(\boldsymbol{\gamma}^\pi\left(c\right))_{c\in\mathcal{C}}$ are such that

$$\begin{aligned}\pi\left(c\right)\left(U_i^E\right)'\left(E\left[\rho^\pi_{C^\pi,i}\right]-U_i^V\left(\mathrm{Var}\left(\rho^\pi_{C^\pi,i}\right)\right)\right)\\ \left(1-2\left(U_i^V\right)'\left(\mathrm{Var}\left(\rho^\pi_{C^\pi,i}\right)\right)\left(\rho^\pi_{c,i}-E\left[\rho^\pi_{C^\pi,i}\right]\right)\right)\\ -\mu^\pi\left(c\right)c_i'\left(\boldsymbol{\rho}_c^\pi\right)+\gamma_i^\pi\left(c\right)=0, \qquad (2.30)\end{aligned}$$

$$\mu^\pi\left(c\right)c\left(\boldsymbol{\rho}_c^\pi\right)=0, \qquad (2.31)$$

$$\gamma_i^\pi\left(c\right)\rho^\pi_{c,i}=0, \qquad (2.32)$$

where c_i' denotes the ith component of the gradient ∇c of the constraint function $c\in\mathcal{C}$.

In developing the above KKT conditions, we used the fact that for any $c\in\mathcal{C}$ and $i\in\mathcal{N}$, $\frac{\partial \mathrm{Var}\left(\rho^\pi_{C^\pi,i}\right)}{\partial\rho_{c,i}}=2\pi(c)\left(\rho^\pi_{c,i}-E\left[\rho^\pi_{C^\pi,i}\right]\right)$.

Next, we find relationships between the optimal solution $(\boldsymbol{\rho}_c^\pi)_{c\in\mathcal{C}}$ of OPT-STAT and OPTAVR. To that end, let $\boldsymbol{\theta}^\pi := (\mathbf{m}^\pi, \mathbf{v}^\pi)$ where for each $i \in \mathcal{N}$, we define

$$
\begin{aligned}
m_i^\pi &:= E\left[\rho_{C^\pi,i}^\pi\right], && (2.33)\\
v_i^\pi &:= \mathrm{Var}^\pi\left(\rho_{C^\pi,i}^\pi\right), && (2.34)\\
e_i^\pi &:= m_i^\pi - U_i^V\left(v_i^\pi\right). && (2.35)
\end{aligned}
$$

Definition 2.4. Let $\mathcal{H}^*$ be the set of fixed points defined by

$$\mathcal{H}^* = \{(\mathbf{m}, \boldsymbol{v}) \in \mathcal{H} : (\mathbf{m}, \boldsymbol{v}) \text{ satisfies } (2.36) - (2.37)\},$$

where

$$E\left[r_i^*\left((\mathbf{m}, \boldsymbol{v}), C^\pi\right)\right] = m_i \qquad \forall\, i \in \mathcal{N}, \tag{2.36}$$

$$\mathrm{Var}\left(r_i^*\left((\mathbf{m}, \boldsymbol{v}), C^\pi\right)\right) = v_i \qquad \forall\, i \in \mathcal{N}. \tag{2.37}$$

Recall that $\mathbf{r}^*\left((\mathbf{m}, \boldsymbol{v}), c\right)$ denotes the optimal solution to OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$ and $\mathcal{H}$ is defined in (2.16). Thus, $\mathcal{H}^*$ is the set of parameter values $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{v})$ that can be viewed as fixed points for 'stationary modification' of AVR obtained by replacing $r_i^*(t)$ and $(r_i^*(t) - m_i(t))^2$ in (2.19) and (2.20) with their expected values. Theorem 2.3 below shows that in fact there is but one such fixed point $\boldsymbol{\theta}^\pi$. A proof is given in Appendix 2.4.1.

Theorem 2.3. *$\boldsymbol{\theta}^\pi$ satisfies the following:*
(a) $\mathbf{r}^*\left(\boldsymbol{\theta}^\pi, c\right) = \boldsymbol{\rho}_c^\pi$ for each $c \in \mathcal{C}$, and
(b) $\mathcal{H}^* = \{\boldsymbol{\theta}^\pi\}$.

Using these results we will study a differential equation that mimics the evolution of the parameters under AVR and show that it converges to $\boldsymbol{\theta}^\pi$.

2.5.2 Convergence of auxiliary ODE associated with AVR

In this subsection, we study and establish convergence of an auxiliary ODE which evolves according to the average dynamics of AVR. This will subsequently be used in establishing convergence properties of AVR.

Consider the following differential equation

$$\frac{d\boldsymbol{\theta}^A(\tau)}{d\tau} = \bar{\mathbf{g}}\left(\boldsymbol{\theta}^A(\tau)\right) + \mathbf{z}\left(\boldsymbol{\theta}^A(\tau)\right), \tag{2.38}$$

for $\tau \geq 0$ with $\boldsymbol{\theta}^A(0) \in \mathcal{H}$ where $\bar{\mathbf{g}}(\boldsymbol{\theta})$ is a function taking values in $\mathbb{R}^{2N}$ defined as follows: for $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{v}) \in \mathcal{H}$, let

$$(\bar{\mathbf{g}}(\boldsymbol{\theta}))_i := E\left[r_i^*(\boldsymbol{\theta}, C^\pi)\right] - m_i, \tag{2.39}$$

$$(\bar{\mathbf{g}}(\boldsymbol{\theta}))_{N+i} := E\left[(r_i^*(\boldsymbol{\theta}, C^\pi) - m_i)^2\right] - v_i. \tag{2.40}$$

In (2.38), $z(\boldsymbol{\theta}) \in -C_{\mathcal{H}}(\boldsymbol{\theta})$ is a projection term corresponding to the smallest vector that ensures that the solution remains in $\mathcal{H}$ (see Section 4.3 of [29]). The set $C_{\mathcal{H}}(\boldsymbol{\theta})$ contains only the zero element when $\boldsymbol{\theta}$ is in the interior of $\mathcal{H}$, and for $\boldsymbol{\theta}$ on the boundary of the set $\mathcal{H}$, $C_{\mathcal{H}}(\boldsymbol{\theta})$ is the convex cone generated by the outer normals at $\boldsymbol{\theta}$ of the faces of $\mathcal{H}$ on which $\boldsymbol{\theta}$ lies. The motivation for studying the above differential equation should be partly clear by comparing the right hand side of (2.38) (see (2.39)-(2.40)) with AVR's update equations (2.19)-(2.20), and we can associate the term $z(\boldsymbol{\theta})$ with the constrained nature of AVR's update equations. The following result shows that $z(\boldsymbol{\theta})$ appearing in (2.38) is innocuous in the sense that we can ignore it when we study the differential equation. The proof shows the redundancy of the term $z(\boldsymbol{\theta})$ by arguing that the differential equation itself ensures that $\boldsymbol{\theta}^A(\tau)$ stays within $\mathcal{H}$.

Lemma 2.4. *For any* $\boldsymbol{\theta} \in \mathcal{H}$, $z_j(\boldsymbol{\theta}) = 0$ *for all* $1 \leq j \leq 2N$.

Proof. Recall that $\mathcal{H} = [0, r_{\max}]^N \times [0, v_{\max}]^N$ and $v_{\max} = r_{\max}^2$. Note that for any $\boldsymbol{\theta}$ in the interior of $\mathcal{H}$, $z_j(\boldsymbol{\theta}) = 0$ for all j such that $1 \leq j \leq 2N$ from the definition of $C_{\mathcal{H}}(\boldsymbol{\theta})$ and thus we can restrict our attention to the boundary of $\mathcal{H}$. For any $\boldsymbol{\theta}$ on the boundary of $\mathcal{H}$ and $i \in \mathcal{N}$, we can use the facts that $(\bar{\mathbf{g}}(\boldsymbol{\theta}))_i = E[r_i^*(\boldsymbol{\theta}, C^\pi)] - m_i$ and $0 \leq r_i^*(\boldsymbol{\theta}, C^\pi), m_i \leq r_{\max}$, to conclude that $z_i(\boldsymbol{\theta}) = 0$. Similarly, since $v_{\max} = r_{\max}^2$, we can show that $z_j(\boldsymbol{\theta}) = 0$ for any j such that $N+1 \leq j \leq 2N$. $\square$

Note that (2.38) has a unique solution for a given initialization due to Lipschitz continuity results in Lemma 2.2.

We define the set $\widetilde{\mathcal{H}} \subset \mathcal{H}$ as follows:

$$\widetilde{\mathcal{H}} := \Big\{ (\mathbf{m}, \boldsymbol{v}) \in \mathcal{H} : \text{ there exists } (\boldsymbol{\rho}_c)_{c \in \mathcal{C}} \in \mathcal{R}_{\mathcal{C}} \text{ such that} \\ E[\rho_{C^\pi, i}] = m_i, \ \mathrm{Var}(\rho_{C^\pi, i}) \leq v_i \leq r_{\max}^2 \quad \forall\, i \in \mathcal{N} \Big\}$$

where $\mathcal{R}_{\mathcal{C}}$ is the set of feasible stationary reward allocation policies defined in (2.29). We can view $\widetilde{\mathcal{H}}$ as the set of all 'achievable' mean variance pairs, i.e., for any $(\mathbf{m}, \boldsymbol{v}) \in \mathcal{H}$ there is some stationary allocation policy with associated mean vector equal to $\mathbf{m}$ and associated variance vector componentwise less than or equal to $\boldsymbol{v}$. Here, the restriction $v_i \leq r_{\max}^2$ for each i ensures that $\widetilde{\mathcal{H}}$ is bounded. Further, for any $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{v}) \in \widetilde{\mathcal{H}}$, let

$$\widetilde{\mathcal{R}}(\boldsymbol{\theta}) \quad := \quad \Big\{ (\boldsymbol{\rho}_c)_{c \in \mathcal{C}} \in \mathcal{R}_{\mathcal{C}} : E[\rho_{C^\pi, i}] = m_i, \ \mathrm{Var}(\rho_{C^\pi, i}) \leq v_i \ \forall\, i \in \mathcal{N} \Big\}.$$

We can view $\widetilde{\mathcal{R}}(\boldsymbol{\theta})$ as the set of all feasible stationary reward allocation policies corresponding to an achievable $\boldsymbol{\theta} \in \widetilde{\mathcal{H}}$.

The following result characterizes several useful properties of the sets introduced above.

Lemma 2.5. *(a) For any $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{v}) \in \widetilde{\mathcal{H}}$, $\widetilde{\mathcal{R}}(\boldsymbol{\theta})$ is a non-empty compact subset of*

$\mathbb{R}^{N|\mathcal{C}|}$.

(b) $\widetilde{\mathcal{H}}$ *is a bounded, closed and convex set.*

Proof. For any $\boldsymbol{\theta} \in \widetilde{\mathcal{H}}$, using the definition of $\widetilde{\mathcal{H}}$, we see that $\widetilde{\mathcal{R}}(\boldsymbol{\theta})$ is a non-empty set. For any $c \in \mathcal{C}$, the set $\left\{\boldsymbol{\rho}_c \in \mathbb{R}^N : c(\boldsymbol{\rho}_c) \leq 0,\ \rho_{c,i} \geq 0 \quad \forall\, i \in \mathcal{N}\right\}$ is compact due to continuity (see Assumption C.1) and boundedness (see Assumption C.2) of feasible region associated with functions in $\mathcal{C}$. Thus, $\mathcal{R}_{\mathcal{C}}$ is also compact. Now, note that $\widetilde{\mathcal{R}}(\boldsymbol{\theta})$ is the intersection of a compact set $\mathcal{R}_{\mathcal{C}}$, and Cartesian product of intersection of inverse images of closed sets associated with continuous functions (corresponding to $\mathbb{E}[.]$ and $\mathrm{Var}(\cdot)$) defined over $\mathbb{R}^N$. Thus, $\widetilde{\mathcal{R}}(\boldsymbol{\theta})$ is compact, and this proves (a).

$\widetilde{\mathcal{H}}$ is bounded since $0 \leq m_i \leq r_{\max}$ and $0 \leq v_i \leq r_{\max}^2$ for each $i \in \mathcal{N}$, and each $(\mathbf{m}, \boldsymbol{v}) \in \widetilde{\mathcal{H}}$.

Let $(\overline{\mathbf{m}}, \overline{\boldsymbol{v}})$ be any limit point of $\widetilde{\mathcal{H}}$. Then, there is a sequence $((\mathbf{m}_n, \boldsymbol{v}_n))_{n\in\mathbb{N}} \subset \widetilde{\mathcal{H}}$, such that $\lim_{n\to\infty}(\mathbf{m}_n, \boldsymbol{v}_n) = (\overline{\mathbf{m}}, \overline{\boldsymbol{v}})$. Let $(\boldsymbol{\rho}_{c,n})_{c\in\mathcal{C}} \in \widetilde{\mathcal{R}}((\mathbf{m}_n, \boldsymbol{v}_n))$ for each $n \in \mathbb{N}$. Since $\left((\boldsymbol{\rho}_{c,n})_{c\in\mathcal{C}}\right)_{n\in\mathbb{N}}$ is a sequence in the compact set $\mathcal{R}_{\mathcal{C}}$, it has some convergent subsequence $\left((\boldsymbol{\rho}_{c,n_k})_{c\in\mathcal{C}}\right)_{k\in\mathbb{N}}$. Suppose that the subsequence converges to $(\overline{\boldsymbol{\rho}}_c)_{c\in\mathcal{C}} \in \mathcal{R}_{\mathcal{C}}$. Then,

$$E\left[\overline{\rho}_{C^\pi,i}\right] = \lim_{k\to\infty} E\left[\rho_{C^\pi,n_k i}\right] = \lim_{k\to\infty} m_{n_k i} = \overline{m}_i,$$
$$\mathrm{Var}\left(\overline{\rho}_{C^\pi,i}\right) = \lim_{k\to\infty} \mathrm{Var}\left(\rho_{C^\pi,n_k i}\right) \leq \lim_{k\to\infty} v_{n_k i} = \overline{v}_i.$$

Thus, $(\overline{\boldsymbol{\rho}}_c)_{c\in\mathcal{C}} \in \widetilde{\mathcal{R}}((\overline{\mathbf{m}}, \overline{\boldsymbol{v}}))$, and hence, $(\overline{\mathbf{m}}, \overline{\boldsymbol{v}}) \in \widetilde{\mathcal{H}}$. Thus, $\widetilde{\mathcal{H}}$ contains all its limit points and hence is closed.

To show convexity, consider $(\mathbf{m}_1, \boldsymbol{v}_1), (\mathbf{m}_2, \boldsymbol{v}_2) \in \widetilde{\mathcal{H}}$, and we show that for any given $\alpha \in [0,1]$, we have $\alpha(\mathbf{m}_1, \boldsymbol{v}_1) + (1-\alpha)(\mathbf{m}_2, \boldsymbol{v}_2) \in \widetilde{\mathcal{H}}$. Let $(\boldsymbol{\rho}_{c,1})_{c\in\mathcal{C}} \in \widetilde{\mathcal{R}}((\mathbf{m}_1, \boldsymbol{v}_1))$ and $(\boldsymbol{\rho}_{c,2})_{c\in\mathcal{C}} \in \widetilde{\mathcal{R}}((\mathbf{m}_2, \boldsymbol{v}_2))$. Hence for each $i \in \mathcal{N}$, $\mathrm{Var}(r_{1i}(C^\pi)) \leq$

v_{1i} and $\mathrm{Var}\left(r_{2i}\left(C^{\pi}\right)\right) \leq v_{2i}$. Let $\boldsymbol{\rho}_{c,3} = \alpha \boldsymbol{\rho}_{c,1} + (1-\alpha)\, \boldsymbol{\rho}_{c,2}$. Thus, for each $i \in \mathcal{N}$,

$$E\left[\rho_{C^{\pi},3i}\right] = \alpha m_1 + (1-\alpha)\, m_2. \tag{2.41}$$

Next, note that $\mathrm{Var}\left(\rho_{C^{\pi}}\right)$ is a convex function of $\left(\rho_c\right)_{c\in\mathcal{C}}$. This can be shown using convexity of square function and linearity of expectation. Thus, for each $i \in \mathcal{N}$,

$$\begin{aligned} \mathrm{Var}\left(\rho_{C^{\pi},3i}\right) &\leq \alpha \mathrm{Var}\left(\rho_{C^{\pi},1i}\right) + (1-\alpha)\, \mathrm{Var}\left(\rho_{C^{\pi},2i}\right) \\ &\leq \alpha v_{1i} + (1-\alpha)\, v_{2i}. \end{aligned} \tag{2.42}$$

From (2.41) and (2.42), we have $\left(\mathbf{r}_3(c)\right)_{c\in\mathcal{C}} \in \widetilde{\mathcal{R}}\left(\alpha\left(\mathbf{m}_1, \boldsymbol{v}_1\right) + (1-\alpha)\left(\mathbf{m}_2, \boldsymbol{v}_2\right)\right)$, and thus $\alpha\left(\mathbf{m}_1, \boldsymbol{v}_1\right) + (1-\alpha)\left(\mathbf{m}_2, \boldsymbol{v}_2\right) \in \widetilde{\mathcal{H}}$. □

The next result gives a set of sufficient conditions to establish asymptotic stability of a point with respect to an ordinary differential equation. This result is a generalization of Theorem 4 in [49].

Lemma 2.6. *Consider a differential equation*

$$\dot{x} = f(x)\,, \qquad x \in \mathbb{R}^d\,, \tag{2.43}$$

where f is locally Lipschitz and all trajectories exist for $t \in [0, \infty)$. Suppose that some compact set $K \subset \mathbb{R}^d$ is asymptotically stable with respect to (2.43) *and also suppose that there exists a continuously differentiable function $L\colon \mathbb{R}^d \to \mathbb{R}$ and some $x_0 \in K$ such that*

$$\nabla L(x) \cdot f(x) < 0 \qquad \forall x \in K\,,\ x \neq x_0\,. \tag{2.44}$$

Then x_0 is an asymptotically stable equilibrium for (2.43) *in $\mathbb{R}^d$.*

Proof. The approach used here is similar to that in [49]. Let $\delta > 0$ be given. With

$B_\delta(x_0)$ denoting the open ball of radius δ centered at x_0 select $\varepsilon \in (0, \delta)$ such that

$$\max_{\overline{B_\varepsilon(x_0)}} L < \min_{K \setminus B_\delta(x_0)} L\,. \tag{2.45}$$

This is possible, since the hypotheses imply that $L(x_0) < L(x)$ for all $x \in K$, $x \neq x_0$. Indeed, consider any solution γ of (2.43) starting at $x \in K$, with $x \neq x_0$. Then the invariance of K and (2.44) imply that the set of ω-limit points of γ is necessarily the singleton $\{x_0\}$. Note that L is non-increasing along trajectories in K and is strictly decreasing along any portion of a trajectory which does not contain x_0. Choose any $t' > 0$ such $\gamma(t) \neq x_0$ for all $t \in [0, t']$ (this is of course possibly by the continuity of $t \mapsto \gamma(t)$). Therefore we must have

$$L(x) = L(\gamma(0)) > L(\gamma(t')) \geq \lim_{t\to\infty} L(\gamma(t)) = L(x_0)\,.$$

Since K is asymptotically stable there exists a decreasing sequence of open sets $\{G_k\}_{k\in\mathbb{N}}$ such that each G_k is invariant with respect to (2.43) and $\cap_{k\in\mathbb{N}} G_k = K$. By (2.44)–(2.45) and the continuity of L and $\nabla L \cdot f$ we can select $n \in \mathbb{N}$ large enough such that

$$\nabla L(x) \cdot f(x) < 0 \qquad \forall x \in \bar{G}_n \setminus B_\varepsilon(x_0) \tag{2.46a}$$

$$\max_{\overline{B_\varepsilon(x_0)}} L < \min_{\bar{G}_n \setminus B_\delta(x_0)} L\,. \tag{2.46b}$$

It is clear by (2.46a)–(2.46b) that any trajectory starting in $G_n \cap B_\varepsilon(x_0)$ stays in $B_\delta(x_0)$, implying that x_0 is a stable equilibrium. Let γ be any trajectory of (2.43). Asymptotic stability of K implies that there exists $t_1 > 0$ such that $\gamma(t) \in G_n$ for all $t > t_1$. Also by (2.46a) there exists $t_2 \geq t_1$ such that $\gamma(t_2) \in G_n \cap B_\delta(x_0)$. Therefore x_0 is asymptotically stable. □

We are now in a position to establish the convergence result for the ODE in

(2.38). The proof relies on the optimality properties of the solutions to OPTAVR, Lemma 3 from [49], Theorem 2.3 (b), and Lemma 2.6.

Theorem 2.4. *Suppose* $\boldsymbol{\theta}^A(\tau)$ *evolves according to the ODE in* (2.38)*. Then, for any initial condition* $\boldsymbol{\theta}^A(0) \in \mathcal{H}$*,* $\lim_{\tau\to\infty} \boldsymbol{\theta}^A(\tau) = \boldsymbol{\theta}^\pi$*.*

Proof. Applying Lemma 3 in [49] and by identifying $\mathcal{V} \equiv \tilde{\mathcal{H}}$, it follows that $\tilde{\mathcal{H}}$ is asymptotically stable for (32). Define

$$L(\boldsymbol{\theta}) = L(\boldsymbol{m}, \boldsymbol{v}) := -\sum_{i\in\mathcal{N}} U_i^E\big(m_i - U_i^V(v_i)\big).$$

Then

$$\begin{aligned} \nabla L(\boldsymbol{\theta}) \cdot \bar{g}(\boldsymbol{\theta}) \quad = \quad & -\sum_{i\in\mathcal{N}} (U_i^E)'\big(m_i - U_i^V(v_i)\big)\Big(\mathbb{E}\big[r_i^*(\boldsymbol{\theta}, C^\pi)\big] - m_i \\ & -(U_i^V)'(v_i)\big(\mathbb{E}\big[(r_i^*(\boldsymbol{\theta}, C^\pi) - m_i)^2\big] - v_i\big)\Big). \end{aligned} \tag{2.47}$$

If $\boldsymbol{\theta} \in \tilde{\mathcal{H}}$, then for some $\boldsymbol{\rho} \in \tilde{\mathcal{R}}(\boldsymbol{\theta})$, (2.47) takes the form

$$\begin{aligned} \nabla L(\boldsymbol{\theta}) \cdot \bar{g}(\boldsymbol{\theta}) = -\mathbb{E}\left[\Phi_{\boldsymbol{\theta}}(\boldsymbol{r}^*(\boldsymbol{\theta}, C^\pi)) - \Phi_{\boldsymbol{\theta}}(\boldsymbol{\rho}_{C^\pi})\right] \\ -\sum_{i\in\mathcal{N}} (U_i^E)'\big(m_i - U_i^V(v_i)\big)\big(U_i^V\big)'(v_i)\big(v_i - \mathrm{Var}\big(\rho_{C^\pi,i}\big)\big) \end{aligned} \tag{2.48}$$

where $\Phi_{\boldsymbol{\theta}}$ is defined in (2.24). The optimality of $r_i^*(\boldsymbol{\theta}, c)$ for OPTAVR$((\boldsymbol{m}, \boldsymbol{v}), c)$ and the fact that $\boldsymbol{\rho} \in \tilde{\mathcal{R}}(\boldsymbol{\theta})$ together with Assumptions U.V.1 and U.E. then imply that both terms on the right-hand-side of (2.48) are nonpositive and that they vanish only if

$$\mathbb{E}\left[r_i^*(\boldsymbol{\theta}, C^\pi)\right] = \mathbb{E}\left[\rho_{C^\pi,i}\right] = m_i, \tag{2.49}$$

$$\mathrm{Var}\big(r_i^*(\boldsymbol{\theta}, C^\pi)\big) = \mathrm{Var}\big(\rho_{C^\pi,i}\big) = v_i\,. \tag{2.50}$$

In turn, by Theorem 3 these imply that $\boldsymbol{\theta} = \boldsymbol{\theta}^{\pi}$. Therefore $\nabla L(\boldsymbol{\theta}) \cdot \bar{g}(\boldsymbol{\theta}) < 0$ for all $\boldsymbol{\theta} \in \tilde{\mathcal{H}}$, $\boldsymbol{\theta} \neq \boldsymbol{\theta}^{\pi}$ and the result follows by Lemmas 4 and 6. □

If the Lipschitz hypothesis in Assumptions U.V.1 and U.E. is relaxed, then the conclusions of Lemma 2.2 hold with continuity replacing Lipschitz continuity. Existence of solutions to the ordinary differential equation (2.38) in the set $\mathcal{H}$ follows by Peano's theorem since $\mathcal{H}$ is compact, thus rendering the vector field (associated with (2.38)) continuous and bounded. Note that Lemma 2.6 does not require Lipschitz continuity, and nor does the proof of Theorem 2.4.

2.5.3 Convergence of AVR and proof of Theorem 2.1

In this subsection, we complete the proof of Theorem 2.1. We first establish a convergence result for the sequence of iterates of the AVR algorithm $(\boldsymbol{\theta}(t))_{t\in\mathbb{N}}$ based on the associated ODE (2.38). We do so by viewing (2.19)-(2.20) as a stochastic approximation update equation, and use a result from [29] that relates the iterates to the ODE (2.38). We establish the desired convergence result by utilizing the corresponding result obtained for the ODE in Theorem 2.4.

Lemma 2.7. *If $\boldsymbol{\theta}(0) \in \mathcal{H}$, then the sequence $(\boldsymbol{\theta}(t))_{t\in\mathbb{N}}$ generated by the Algorithm AVR converges almost surely to $\boldsymbol{\theta}^{\pi}$.*

Proof. This proof draws on standard techniques from stochastic approximation (see e.g., [29]). The key idea is to view (2.19)-(2.20) as a stochastic approximation update equation, and using Theorem 1.1 of Chapter 6 from [29] to relate (2.19)-(2.20) to the ODE (2.38). Below, for brevity, we provide details drawing heavily on the framework developed in [29].

In the following, we show that all the Assumptions required to use the theorem are satisfied. The following sets, variables and functions H, $\boldsymbol{\theta}_t$, $\boldsymbol{\xi}_t$, $\mathbf{Y}_t$, ϵ_t, sigma algebras $\mathcal{F}_t$, $\boldsymbol{\beta}_t$, $\delta\mathbf{M}_t$ and the function $\mathbf{g}$ appearing in the exposition of Theorem 1.1

of [29], correspond to the following variables and functions in our problem setting: $H = \mathcal{H}$, $\boldsymbol{\theta}_t = (\mathbf{m}(t), \boldsymbol{v}(t))$, $\xi_t = c_t$, for each $i \in \mathcal{N}$ $(Y_t)_i = r_i^*(t) - m_i(t)$ and $(Y_t)_{i+N} = (r_i^*(t) - m_i(t))^2 - v_i(t)$, $\epsilon_t = \frac{1}{t}$ for each t, $\mathcal{F}_t$ is such that $(\boldsymbol{\theta}_0, \mathbf{Y}_{i-1}, \xi_i, i \leq t)$ is $\mathcal{F}_t$-measurable, $\boldsymbol{\beta}_t = \mathbf{0}$ and $\delta \mathbf{M}_t = \mathbf{0}$ for each t, $(g((\mathbf{m}, \boldsymbol{v}), c))_i = r_i^*((\mathbf{m}, \boldsymbol{v}), c) - m_i$ and $(g((\mathbf{m}, \boldsymbol{v}), c))_{i+N} = (r_i^*((\mathbf{m}, \boldsymbol{v}), c) - m_i)^2 - v_i$,

Equation (5.1.1) in [29] is satisfied due to our choice of ϵ_t, and (A4.3.1) is satisfied due to our choice of $\mathcal{H}$. Further, (A.1.1) is satisfied as the solutions to OPTAVR are bounded. (A.1.2) holds due to the continuity result in Lemma 2.2 (a).

We next show that (A.1.3) holds by choosing the function $\bar{\mathbf{g}}$ as follows for each $i \in \mathcal{N}$:

$$
\begin{aligned}
(\bar{g}(\mathbf{m}, \boldsymbol{v}))_i &= E\left[r_i^*((\mathbf{m}, \boldsymbol{v}), C^\pi)\right] - m_i, \\
(\bar{g}(\mathbf{m}, \boldsymbol{v}))_{i+N} &= E\left[(r_i^*((\mathbf{m}, \boldsymbol{v}), C^\pi) - m_i)^2\right] - v_i.
\end{aligned}
$$

Note that the continuity of the function $\bar{\mathbf{g}}$ follows from Lemma 2.2 (b).

From Section 6.2 of [29], if ϵ_t does not go to zero faster than the order of $\frac{1}{\sqrt{t}}$, for (A.1.3) to hold, we only need to show that the strong law of large numbers holds for $(g(\mathbf{m}, \boldsymbol{v}, C_t))_t$ for any $\widehat{\mathbf{q}}$. The strong law of large numbers holds since $(C_t)_{t \in \mathbb{N}}$ is a stationary ergodic random process and g is a bounded function. Assumptions (A.1.4) and (A.1.5) hold since $\boldsymbol{\beta}_t = \mathbf{0}$ and $\delta \mathbf{M}_t = \mathbf{0}$ for each t. To check (A.1.6) and (A.1.7), we use sufficient conditions discussed in [29] following Theorem 1.1. Assumption (A.1.6) holds since g is bounded. (A.1.7) holds due to the continuity of $g((\mathbf{m}, \boldsymbol{v}), c)$ in $(\mathbf{m}, \boldsymbol{v})$ uniformly in c which follows from the continuity result in Lemma 2.2 (a), and the finiteness of $\mathcal{C}$. Thus, using Theorem 1.1, we can conclude that on almost all sample paths, $(\boldsymbol{\theta}(t))_{t \in \mathbb{N}}$ converges to some limit set of the ODE (2.38) in $\mathcal{H}$. From Theorem 2.4, for any initialization in $\mathcal{H}$, this limit set is the singleton $\{\boldsymbol{\theta}^\pi\}$, and thus the main result follows. □

If we use AVR with a constant step size stochastic approximation algorithm obtained by replacing $1/t$ in (2.19)-(2.20) with a small positive constant ϵ, we can use results like Theorem 2.2 from Chapter 8 of [29] to obtain a result similar in flavor to that in Lemma 2.7 (which can then be used to obtain optimality results).

Now we prove Theorem 2.1 mainly using Lemma 2.7, and stationarity and ergodicity assumptions. For each $i \in \mathcal{N}$,

$$\frac{1}{T}\sum_{t=1}^{T} r_i^*(\boldsymbol{\theta}(t), C_t) = \frac{1}{T}\sum_{t=1}^{T}\big(r_i^*(\boldsymbol{\theta}(t), C_t) - r_i^*(\boldsymbol{\theta}^\pi, C_t)\big) + \frac{1}{T}\sum_{t=1}^{T} r_i^*(\boldsymbol{\theta}^\pi, C_t)\,. \tag{2.51}$$

The first term of (2.51) converges to 0 a.s. (i.e., for almost all sample paths) as $T \to \infty$ by Lemma 2.7, the continuity of $\mathbf{r}^*(\boldsymbol{\theta}, c)$ in $\boldsymbol{\theta}$ (see Lemma 2.2 (a)) and the Dominated Convergence Theorem (see, for e.g., [22]). The second term converges to $\mathbb{E}\big[r_i^*\big(\boldsymbol{\theta}^\pi, C^\pi\big)\big]$ by Birkhoff's Ergodic Theorem (see, for e.g., [21]). Now, note that $\mathbb{E}\big[r_i^*\big(\boldsymbol{\theta}^\pi, C^\pi\big)\big] = m_i^\pi$ (see Theorem 2.3 (b) and (2.36)). Since by Lemma 2.7, $\lim_{t\to\infty} m_i(t) = m_i^\pi$, part (a) of Theorem 2.1 is proved.

Next, we prove part (b). Note that for each $i \in \mathcal{N}$,

$$\begin{aligned}
\mathrm{Var}^T(r_i^*) &= \frac{1}{T}\sum_{t=1}^{T}\left(r_i^*\big(\boldsymbol{\theta}(t), C_t\big) - \frac{1}{T}\sum_{s=1}^{T} r_i^*\big(\boldsymbol{\theta}(s), C_s\big)\right)^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}\big(r_i^*\big(\boldsymbol{\theta}(t), C_t\big) - m_i^\pi\big)^2 - \left(\frac{1}{T}\sum_{s=1}^{T} r_i^*\big(\boldsymbol{\theta}(s), C_s\big) - m_i^\pi\right)^2.
\end{aligned} \tag{2.52}$$

The second term on the right-hand-side of (2.52) converges a.s. to zero as $t \to \infty$ by part (a). Also, following the same steps as in the proof of part (a), we see that the first term converges a.s. to v_i^π as $T \to \infty$. Since by Lemma 2.7, $\lim_{t\to\infty} v_i(t) = v_i^\pi$, part (b) of Theorem 2.1 is proved.

Part (c) of Theorem 2.1 follows from parts (a) and (b).

2.6 Simulations

In this section, we evaluate additional performance characteristics of AVR via simulation. We focus on the realization of different mean-variability-fairness tradeoffs by varying the functions $\left(U_i^E, U_i^V\right)_{i\in\mathcal{N}}$, and on the convergence rate of the algorithm.

For the simulations, we consider a time-slotted setting involving time varying utility functions as discussed in Section 2.2. We consider a network where $N = 20$. Temporal variations in video content get translated into time varying quality rate maps, and we model this as follows: in each time slot, a time varying quality rate map for each user is picked independently and uniformly from a set $\mathcal{Q} = \{q_1, q_2\}$. Motivated by the video distortion versus rate model proposed in [51], we consider the following two (increasing) functions that map video compression rate w to video quality

$$q_1(w) = 100 - \frac{40000}{w - 500},\ \ q_2(w) = 100 - \frac{80000}{w - 500}.$$

These (increasing) functions map video compression rate w to a video quality metric. We see that the map q_2 is associated with a time slot in which the video content (e.g., involving a scene with a lot of detail) is such that it needs higher rates for the same quality (when compared to that for q_1). Referring Section 2.2, we see that $\mathcal{Q}_i = \mathcal{Q}$ for each user $i \in \mathcal{N}$. For each user, the peak data rate in each time slot is modeled as an independent random variable with various distributions (discussed below) from the set $\mathcal{W} = \{\omega_1, \omega_2\}$ where $\omega_1 = 30000$ units and $\omega_2 = 60000$ units (thus $\mathcal{P} = \mathcal{W}^N$). Further, we choose $r_{\max} = 100$ and the run length of each simulation discussed below is 100000 time slots.

To obtain different tradeoffs between mean, variability and fairness, for each $i \in \mathcal{N}$ we set $U_i^E(e) = \frac{e^{1-\alpha}}{1-\alpha}$ and $U_i^V(v) = \beta\sqrt{v+1}$ and vary α and β. For a given α, note that $U_i^E(\cdot)$ corresponds to $\alpha-$fair allocation discussed in Section 2.2 where a larger α corresponds to a more fair allocation of QoE. Also, by choosing a larger β we can impose a higher penalty on variability. The choice of $U_i^V(\cdot)$ roughly corresponds to the metric proposed in [59]. To obtain *a good initialization* for AVR, the reward allocation in the first 10 time slots is obtained by solving a modified version of OPTAVR$((\mathbf{m}, \boldsymbol{v}), c)$ with a simpler objective function $\sum_{i \in \mathcal{N}} U_i^E(r_i)$ (which does not rely on any estimates) under the same constraints (2.17) and (2.18), and run AVR from the 11th time slot initialized with parameters $(\mathbf{m}, \boldsymbol{v})$ set to the mean reward and half the variance in reward over the first ten time slots.

We first study a homogeneous setting in which, for each time slot, the peak data rate of each user is picked independently and uniformly at random from the set $\mathcal{W}$. Here, we set $\alpha = 1.5$ and vary β over $\{0.02, 0.1, 0.2, 0.5, 1, 2\}$. The averaged (across users) values of the mean reward and standard deviation of the reward allocation for the different choices of β are shown in Fig 2.1. Not only does the

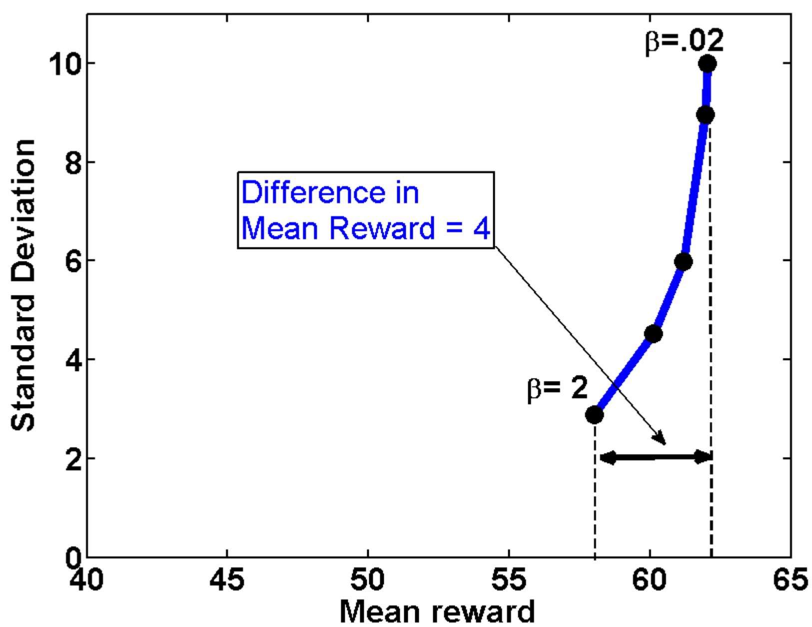

Figure 2.1: Homogeneous setting: Mean-Variability tradeoffs

standard deviation reduce with a higher β, we also see that the reduction in mean reward for a given reduction in variability is very small. For instance here we were able to reduce the standard deviation in reward from around 10 to 3 (i.e., around 70 % reduction) at the cost of a mere reduction of 4 units in the mean reward (around 7 % reduction). It should be clear that the reduction in variance corresponding to the above data will be even more drastic than that of the standard deviation and this is the case in the next setting too.

Next, we study a heterogeneous setting. For each time slot, the peak data rate of each user indexed 1 through 10 is modeled as a random variable taking values ω_1 and ω_2 with probability 0.9 and 0.1 respectively, and that of each user indexed 11 through 20 is ω_1 and ω_2 with probability 0.1 and 0.9 respectively. Thus, in this setting, users with index in the range 1 through 10 typically see poorer channels, and can end up being sidelined if the allocation is not fair. To measure the fairness of a reward allocation, we use a simple metric M_{fair} which is the ratio of the minimum value to the maximum value of the QoE of the users. In Table 2.2, the value of M_{fair} along with values of the averages (across users) of the mean, variance and standard deviation of the allocated rewards for different choices of α and β are given. As in the homogeneous setting, we see that we can achieve drastic reduction in the variability of quality (measured in terms of either the variance or the standard deviation) for a relatively small reduction in the mean reward. We further see that higher values of α result in a higher values of M_{fair} for the same β, and thus reduce the disparity in allocation of quality.

Fig. 2.2 depicts the evolution of the parameters (m_5, v_5) and (m_{15}, v_{15}) (i.e., the parameters associated with Users 5 and 15) for the heterogeneous setting where $\alpha = 1.5$ and $\beta = 0.5$. Also, note that User 5 (depicted using dashed lines) sees poorer channel conditions than User 15. Recall that convergence of the parameters $\mathbf{m}$ and $\mathbf{v}$ is the key property used in establishing optimality in Theorem 2.2. Thus, we can conclude that reward allocations under AVR are close to optimal after a few

Table 2.2: Heterogeneous setting: Mean-Variability-Fairness Tradeoffs

α	β	Mean	Variance	Std.Devn.	M_{fair}
0.05	0.02	62.14	65.23	8.08	0.85
0.05	0.1	62.10	49.11	7.01	0.84
0.05	0.5	61.44	19.52	4.42	0.83
0.05	1	60.72	11.66	3.41	0.81
0.05	2	59.25	5.15	2.27	0.79
1.5	0.02	62.06	65.09	8.07	0.89
1.5	0.1	62.00	49.20	7.01	0.89
1.5	0.5	61.37	19.67	4.43	0.88
1.5	1	60.66	11.87	3.44	0.87
1.5	2	59.21	5.23	2.29	0.86
5	0.02	61.86	65.89	8.10	0.93
5	0.1	61.80	49.72	7.05	0.93
5	0.5	61.18	20.122	4.47	0.93
5	1	60.46	11.80	3.44	0.93
5	2	58.87	5.03	2.24	0.92

hundred time slots by which time the parameters have roughly converged.

In Fig. 2.3, the dashed lines depict the performance of AVR in terms of $\phi_T(\mathbf{r}^*)$ for different simulation runs of the heterogeneous setting where $\alpha = 1.5$ and $\beta = 0.5$. The thick line exhibits the value of $\sum_{i\in\mathcal{N}} U_i^E \left(m_i^\pi - U_i^V(v^\pi)\right)$ with the limiting estimated parameters $\mathbf{m}^\pi$ and $\mathbf{v}^\pi$ obtained by running AVR for 100000 slots. We once again observe good rate of convergence similar to that of the estimated parameters shown Fig. 2.2. Note that $\sum_{i\in\mathcal{N}} U_i^E \left(m_i^\pi - U_i^V(v^\pi)\right)$ is also an upper bound on the performance of optimal offline scheme over T slots, i.e., the optimal value of OPT(T), as T goes to infinity.

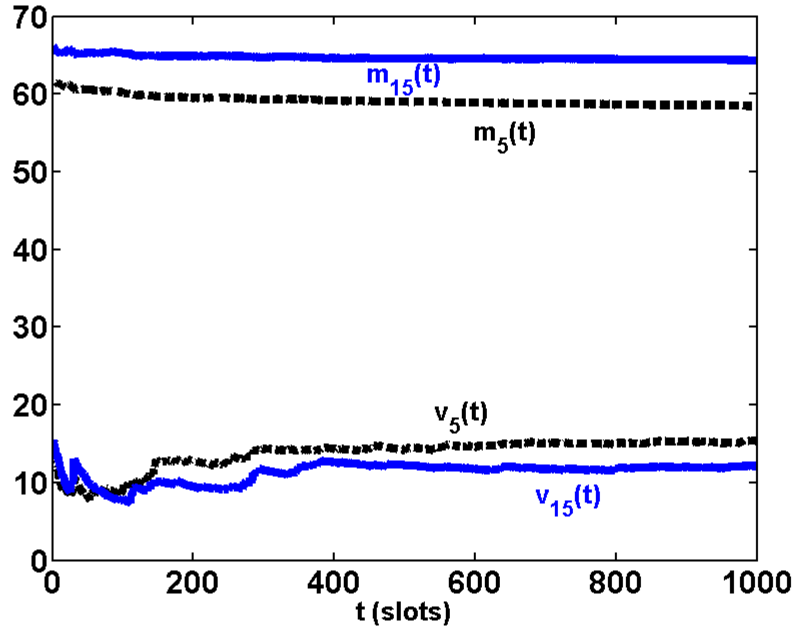

Figure 2.2: Heterogeneous setting: Convergence

2.7 Conclusions

This work presents an important generalization of NUM framework to account for the deleterious impact of temporal variability allowing for tradeoffs between mean, fairness and variability associated with reward allocations across a set of users. We proposed a simple asymptotically optimal online algorithm AVR to solve problems falling in this framework. We believe such extensions to capture variability in reward allocations can be relevant to a fairly wide variety of systems. Our future work will encompass the possibility of addressing resource allocation in systems with buffering or storage. e.g., energy and/or data storage.

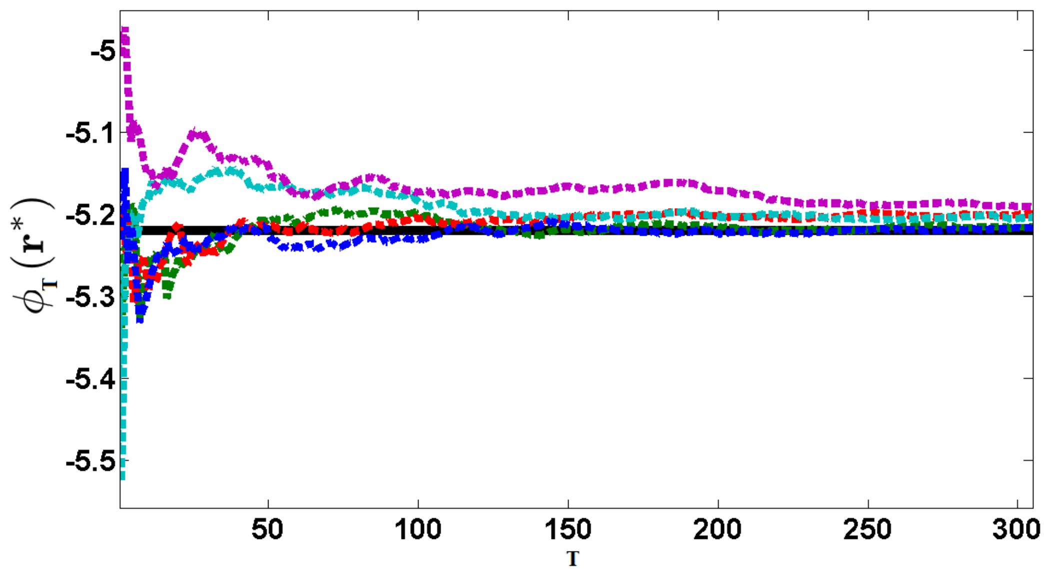

Figure 2.3: Performance of AVR: Evolution of $\phi_T(\mathbf{r}^*)$ for different simulation runs

Chapter 3

NOVA: QoE-driven Optimization of Video Delivery in Networks

3.1 Introduction

There has been tremendous growth in video traffic in the past decade. Current trends (see [12]) suggest that mobile video traffic will more than double each year till 2015, with two-thirds of mobile data traffic being video by 2015. It is unlikely that wireless infrastructure, e.g., base stations, access points, capacity etc, can keep up with such growth. Even densification does not resolve the problem since the variability in throughput is likely to increase or worsen due to increased sensitivity to the dynamic number of users sharing an access point and/or dynamic interference. Given these challenges, optimizing video delivery to make the best use of available network resources is one of the most important networking problems today.

The main focus of this chapter is to develop solutions for optimizing the delivery of stored video, i.e., video stored in video servers, streamed by video clients that can *adapt* their video quality. Our solution is designed to achieve better Quality of Experience (QoE) while taking important video client preferences like rebuffering and data cost into account. Further, it is suited for operation in settings that present video clients with heterogeneous preferences, channels and video content.

In this chapter, we view the video delivery optimization problem for a network as a problem of *fairly* maximizing the video clients' QoE subject to network constraints. Here, QoE is a proxy for 'video client satisfaction'. A comprehensive solution to this problem requires two components- an allocation component and an adaptation component. The allocation component decides how the resources (e.g., bandwidth, power etc) in the network are allocated to the video clients. The adaptation component decides how the video clients adapt their video quality (or video compression rate) in response to the allocated resources, the nature of the video etc. In this chapter, we develop a distributed algorithm Network Optimization for Video Adaptation (NOVA) to jointly optimize the two components. The adaptation component itself has strong optimality guarantees, and can be used in standalone video clients and, in particular, the adaptation in NOVA can be used with video clients using the DASH (Dynamic Adaptive Streaming over HTTP) framework. Under the DASH framework, video associated with each video client is stored at the respective video server (at the content provider), and is a concatenation of several short duration videos called segments which for example could be a GOP (Group Of Pictures). Various 'representations' of a segment are obtained by compressing it to different sizes by changing various parameters associated with the segment like quantization, resolution, frame rate etc, and typically high quality representations of a segment are larger in size. Video clients can *adapt* their video quality across segments, i.e., can pick different representations for different segments. The choice of representations can be based on several factors such as the state of playback buffer, current channel capacity, features of video content being downloaded etc.

For instance, the video client can request representations of smaller size to *adapt* to poor channel conditions.

We identify the following four key factors determining the QoE of a video client: (a) average quality, (b) temporal variability in quality, (c) fraction of time spent rebuffering, and (d) cost to the video client and video content provider. Our main focus is on solving the video delivery optimization problem OPT-BASIC given below which takes these key factors into account:

$$\max \quad \sum_{i \in \mathcal{N}} U_i^E \left(\text{Mean Quality}_i - \text{Quality Variability}_i\right) \tag{3.1}$$

$$\text{subject to} \;\; \text{Rebuffering}_i,\ \text{Cost}_i,\ \text{and Network constraints,} \tag{3.2}$$

where $\mathcal{N}$ is the set of video clients supported by the network and U_i^E is a 'nice' concave function chosen in accordance with the fairness desired in the network. Network constraint captures time varying constraints on network resource allocation allowing us to model wide range variability in resource availability found in real networks.

Next, we discuss the four key factors mentioned above. We measure mean quality for a video session as the average across segments of Short Term Quality (STQ) associated with the downloaded representations of (short duration, e.g., 1 second) segments. STQ of a downloaded segment should ideally capture the viewer's *subjective* evaluation of the quality of the downloaded representation of the segment, although in practice, this subjective metric will be measured approximately using *objective* video quality assessment metrics (see [42] for a survey) like PSNR, SSIM, MSSSIM etc (see [56, 57]).

While the benefit of high mean quality is clear, the detrimental impact of temporal variability on QoE (see [59, 28, 40]), and fundamental tradeoff between average quality and temporal variability in quality is often ignored. Indeed [59] even suggests that temporal variability in quality can result in a QoE that is *worse* than that of a constant quality video with *lower* average quality. Two prominent sources for such variability ([24]) are time varying capacity and the time varying nature of video content. Time varying capacity is especially relevant when considering wireless networks where such variations can be caused by fast fading (on faster time scales, e.g., ms) and slow fading due to shadowing, dynamic interference, mobility, and changing loads (on slower time scales, e.g. secs). The second, is the time varying nature of the dependence of a segment's STQ on parameters like compression rate. Perhaps the key contributor to such change is the video content itself, for instance, segments of same size and same duration could have very different STQ, for e.g., consider two such segments where the first segment is of an action scene (where there is a lot of changing visual content) and the second segment is of a slower scene (where things stay the same).

Rebuffering is the event when playback buffer of a video client empties, and video playback stalls. Rebuffering events have a significant impact on QoE. Indeed [35] points out that the total time spent rebuffering and the frequency of rebuffering events during a video session can significantly reduce video QoE. In our approach, we impose constraints on the fraction of the total time spent rebuffering, and suggest simple ideas to reduce the frequency of rebuffering events. We also provide flexibility to the video client in setting these constraints according to their preferences. For instance, a video client who is willing to tolerate rebuffering in return for higher mean quality (for e.g., to watch a movie in HD over a poor network) can set these constraints accordingly. Such constraints driven by video client preferences will often be content and device dependent, and capture important tradeoffs for the video client.

Client preferences concerning the cost of video delivery are also significant, and are important when viewers wish to manage their wireless data costs. Note that content providers may also pay Content Distribution Network operators for the delivery of video data. Thus, if the cost of data delivery is high, higher QoE often comes at higher cost, and the video client/content provider may want to tradeoff QoE versus delivery cost. In our framework, we allow each video client/content provider to set a constraint on the average cost per unit video duration which in turn reflects the desired tradeoff.

3.1.1 Main contributions

The main contribution of this chapter is a general optimization framework for stored video delivery optimization, that factors heterogeneity in client preferences, QoE models, capacity and video content.

We develop a *simple online* algorithm NOVA (Network Optimization for Video Adaptation) to solve the video delivery optimization problem. Key features of NOVA are listed below:

1. *Optimality:* We establish a strong asymptotic optimality result for NOVA which roughly guarantees that NOVA performs as well as the optimal offline scheme which is omniscient, i.e., knows everything about the evolution of channel and video ahead of time.

2. *Simple and Online*: NOVA only utilizes current information, and is computationally light.

3. *Distributed*: NOVA uses minimal signaling, and can be implemented in a distributed manner.

4. NOVA is *asynchronous* and requires almost no statistical information about the system

5. *Optimal Adaptation:* The adaptation proposed in NOVA is independently optimal, and the optimality properties of the adaptation component of NOVA

is 'insensitive' to the resource allocation, i.e., does not depend on detailed characteristics (for e.g., the specific resource allocation algorithm, time scale of operation etc) of the latter. Further, the adaptation proposed in NOVA is entirely client driven and is also optimal for standalone video clients.

6. *Suited for current practical systems:*

   (a) *Suited for DASH:* The adaptation proposed in NOVA is suited for DASH framework as it entirely client driven, and can be carried out in an asynchronous manner.

   (b) The resource allocation proposed in NOVA requires simple modification of legacy schedulers like proportionally fair schedulers, and can be extended for use in the presence of data users.

3.1.2 Related work

The problem of video delivery optimization in wireless networks has been studied in many works, for instance, see [52, 23, 60, 20, 24, 5] etc. In [20], the problem of optimizing network resource allocation for maximizing the discounted sum of the aggregate quality of the users over time is considered, and a scheme based on Markov decision programming is proposed. Video quality adaptation is not considered in [20], and the main focus of the paper is real time interactive video applications which present an additional challenge of meeting strict deadlines associated with video delivery. [23] is another work aimed at similar applications with strict deadlines, and considers the problem of optimizing both network resource allocation and video quality adaptation under additional assumptions which allow decoupling of the two tasks. [52] focuses on a WLAN setting and proposes a scheme that greedily maximizes the minimum quality among users by determining the optimal encoding rate and physical layer parameters so as to minimize the sum of the distortion caused by source compression and the expected distortion resulting from packet loss during transmission. Although [5] considers the problem of video delivery optimization in a more general network setting (specifically, one involving multiple small base stations serving video clients) which presents new challenges (even ordering of segments is important), the paper considers simpler QoE models that ignore the impact of temporal variability. Further, the algorithm proposed in [5] relies on synchronous quality adaptation decisions which limits the algorithm's ability to exploit good channel conditions, for instance, even if segments are downloaded quickly, one has to wait till the end of a slot to request the next segment. The papers [5] and [60] do not explicitly target rebuffering, and try to ensure low rebuffering through stability of video data queues at the basestation. [60] also considers the problem of video delivery in wireless networks, and proposes a solution based on dynamic programming framework which is computationally heavy and requires detailed knowledge of system statistics. A major weakness of [52, 23, 60, 20, 5] is the limited nature of the QoE models considered that are essentially just the mean quality (or in some cases, a mean of a function of quality across segments) which does not explicitly account for the impact of temporal variability.

There are several works, for e.g., [14], that propose schemes to reduce variability in coding rates to reduce the variability in STQ. But, these approaches ignore the (time varying nature of the) dependence of STQ on rates, and hence are suboptimal. The problem of reducing the variability in quality for SVC coded video over the Internet was considered in [28]. However, they restrict their attention to a single video stream, and they only focus on the reduction of switching rate of quality which is a crude metric for variability of STQ.

The problem of realizing optimal mean-variability tradeoffs is carefully studied in [24] for the video delivery problem, and in Chapter 2 for more general resource allocation problems. While [24] presents a novel algorithm for realizing mean-variability tradeoffs for video delivery, the model considered in [24] (and Chapter 2) involves a strong, and to somewhat impractical, assumption of synchrony- the segment downloads of all the video clients are synchronous, i.e., the download of a segment of each video client starts at the beginning of a (network) slot and finishes at the end of the slot. This assumption on synchrony precludes any explicit control over rebuffering at the video clients and forces the solution in [24] to be a centralized one. The relaxation of assumption on synchrony in this chapter allows us to obtain a distributed asynchronous solution in which the video clients and network controllers operate at their own pace. This relaxation also presents new technical challenges in dealing with distributed asynchronous algorithms operating in a stochastic setting, and the rebuffering constraints in our asynchronous setting effectively induce a new type of constraints involving averages measured over two time scales. Further, the framework incorporates novel heterogeneous client preferences on rebuffering and data costs.

Our work in this chapter relies heavily on results from theory of asynchronous stochastic approximation discussed in Chapter 12 of [29]. We also use extensions of several theoretical tools from works [26], [50] etc related to Network Utility Maximization (NUM).

3.1.3 Notation

Here, we describe some of the notation used in this chapter. We shall consider a network shared by a set $\mathcal{N}$ of video clients (or other entities) where $N{:=}|\mathcal{N}|$ denotes the number of video clients in the system. We use bold letters to denote vectors, e.g., $\mathbf{a} = (a_i)_{i\in\mathcal{N}}$. Given a collection of T objects $(a(t))_{1\leq t\leq T}$ or a sequence $(a(t))_{t\in\mathbb{N}}$, we let $(a)_{1:T}$ denote the finite length sequence $(a(t))_{1\leq t\leq T}$ (in the space associated with the objects of the sequence). For example, consider a sequence $(\mathbf{a}(t))_{t\in\mathbb{N}}$ where each element is a vector. Then $(\mathbf{a})_{1:T}$ denotes the T length sequence containing the first

T vectors of the sequence $(\mathbf{a})_{1:T}$, and $(a_i)_{1:T}$ denotes the sequence containing the ith component of the first T vectors. Let $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, real numbers and nonnegative real numbers respectively. For any function U on $\mathbb{R}$, let U' denote its derivative. For any positive integer M, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^M$ and set $\mathcal{A} \subset \mathbb{R}^M$, let

$$d_M(\mathbf{a}, \mathbf{b}) := \sqrt{\sum_{i=1}^{M} (a_i - b_i)^2}, \quad d_M(\mathbf{b}, \mathcal{A}) := \inf_{\mathbf{a} \in \mathcal{A}} d_M(\mathbf{a}, \mathbf{b}).$$

3.1.4 Organization of the chapter

Section 3.2 introduces the system model and assumptions. We formulate the problem OPT-BASIC as an offline optimization problem in Section 3.3. In section 3.4, we present NOVA to essentially solve the offline optimization problem, and discuss its optimality properties. Section 3.5 is devoted to the proof of optimality of NOVA. We discuss few useful extensions of NOVA in Section 3.6, and conclude the chapter in Section 3.7.

3.2 System model

We consider a network serving video to a fixed set of video clients $\mathcal{N}$ where $|\mathcal{N}| = N$. We consider a slotted network system where resources are allocated for the duration of a slot τ_{slot}, and the slots are indexed by $k \in \{0, 1, 2...\}$.

We assume that resource allocation is subject to time varying constraints. In each slot k, a network controller (e.g., base station in a cellular network) allocates $\mathbf{r}_k = (r_{i,k})_{i \in \mathcal{N}} \in \mathbb{R}_+^N$ bits (or $\mathbf{r}_k/\tau_{slot}$ bits per second) to the video clients such that $c_k(\mathbf{r}_k) \leq 0$, where c_k is a real valued function modeling the current constraints on network resource allocation. In many practical settings, the set of feasible resource allocations in a slot may be *discrete* (i.e., we have to pick from a *finite* set of feasible allocations), and we discuss such settings in Section 4.4.1. However, in this Chapter, we consider the set of feasible allocations in slot k to be $\left\{\mathbf{r}_k \in \mathbb{R}_+^N : c_k\left(\mathbf{r}_k\right) \leq 0\right\}$ which is determined by the function c_k. This function could be determined by various parameters like video clients' SNR, interference etc. In the sequel, we refer to these functions as allocation constraints. Let C_k denote the random variable corresponding to the allocation constraint in slot k (and c_k is a realization of it). Even though we are assuming We make the following assumptions on these allocation constraints:

Assumptions C.1-C.3 (Time varying allocation constraints)

C.1 $\left(C_k\right)_{k\in\mathbb{N}}$ is a *stationary ergodic* process of functions selected from a set $\mathcal{C}$.

C.2 $\mathcal{C}$ is a (arbitrarily large) finite set of real valued functions on $\mathbb{R}_+^N$, such that each function $c \in \mathcal{C}$ is *convex* and continuously differentiable on an open set containing $\left[0, r_{\max}\right]^N$ with $c\left(\mathbf{0}\right) \leq 0$ and

$$\min_{\mathbf{r}\in\left[0,r_{\max}\right]^N} c\left(\mathbf{r}\right) < 0. \tag{3.3}$$

C.3 The feasible region for each allocation constraint is *bounded*: there is a constant $0 < r_{\max} < \infty$ such that for any $c \in \mathcal{C}$ and $\mathbf{r} \in \mathbb{R}_+^N$ satisfying $c\left(\mathbf{r}\right) \leq 0$, we have $r_i \leq r_{\max}$ for each $i \in \mathcal{N}$.

As indicated in Assumption C.1, we model the evolution of the allocation constraints as a stationary ergodic process. Hence, time averages associated with the allocation constraints will converge to their respective statistical averages, and the distribution of the random vector $\left(C_{k_1+s}, C_{k_2+s}, ..., C_{k_n+s}\right)$ for any choice of indices $k_1, ..., k_n$ does not depend on the shift s, thus the marginal distribution of C_k does not depend on time. We denote the marginal distribution of this process by $(\pi(c))_{c\in\mathcal{C}}$. Without loss of generality, we assume that $\pi^{\mathcal{C}}(c) > 0$ for each $c \in \mathcal{C}$. This model (along with the generalization in Subsection 3.6.2) captures a fairly general class of allocation constraints, including, for example, time-varying capacity constraints associated with bandwidth allocation in wireless networks.

We express the network constraints in (3.2) of OPT-BASIC as the requirement $c_k(\mathbf{r}_k) \leq 0$ on resource allocation $\mathbf{r}_k$ in each slot k. We impose an additional requirement on the resource allocation algorithm to ensure that the resource allocation to each video client $i \in \mathcal{N}$ in each slot should be at least $r_{i,\min}$ where $r_{i,\min}$ is a small positive constant. This technical requirement can be relaxed as long as we ensure that each video client can be guaranteed a strictly positive amount of resource allocation over a fixed (large) number of slots.

Next, we discuss our video quality adaptation model which is compatible with that proposed in DASH. The video associated with each video client $i \in \mathcal{N}$ is stored at the respective video server (at the content provider), and is a concatenation of segments. Representations of a segment are obtained by compressing it to different sizes by changing various parameters associated with it like quantization, resolution, frame rate etc. Video clients adapt their quality across segments by selecting different representations for different segments, and these choices can be based on a variety of factors such as the state of playback buffer, current channel capacity, features of video content being downloaded etc.

The STQ of a (downloaded) segment, measured using objective video quality assessment metrics like PSNR, SSIM, MSSSIM etc, typically increases with the compression rate of the corresponding downloaded representation. Here the compression rate is the ratio of the size of the segment's representation to the duration of the segment. Note that the size and (hence the) compression rate *also* depend on the size of overheads due to metadata (like identifiers, sequence numbers etc) associated with the generation of the data-unit (e.g., file) associated with the segment's representation. In the sequel, we interchangeably use the terms quality and STQ. We abstract the relationship between the compression rate and quality of a segment using a convex increasing function referred to as QR (Quality Rate) tradeoff. QR tradeoffs maps quality q to the compression rate $f_s(q)$ (measured in bits per second). Note that for each segment and given compression rate, we are implicitly restricting our attention to the representation with highest quality and ignoring less efficient representations. QR tradeoffs can be segment dependent and vary depending on the nature of the segment's video content. For instance, a segment associated with a slow scene (where things stay the same) will typically have a 'steeper' QR tradeoff when compared to that of an action scene (where there is a lot of changing visual content). For stored video these functions might be obtained offline. For video streaming of live events, live broadcast of TV channels etc, computationally efficient video quality assessment metrics can be used to obtain the QR tradeoffs.

Let $f_{i,s}$ denote a realization of QR tradeoff associated with the sth segment downloaded by video client i. Also, let $F_{i,s}$ denote the random variable corresponding to the QR tradeoff associated with the sth segment of video client i. Next, let $l_{i,s}$ denote a realization of length (or duration in seconds) of the sth segment downloaded by video client i, and let $L_{i,s}$ denote the corresponding random variable. Thus, to obtain a quality q for the sth segment, the size of the segment that has to be downloaded by video client i is given by $l_{i,s} f_{i,s}(q)$. For each video client $i \in \mathcal{N}$, we make the following assumptions on the QR tradeoffs and segment lengths associated with it:

Assumptions QRL.1-QRL.3 on QR tradeoffs and segment lengths

QRL.1 $(F_{i,s}, L_{i,s})_{s \geq 0}$ is a *stationary ergodic* process taking values in a set $\mathcal{FL}_i \subset \mathcal{F}_i \times \mathcal{L}_i$.

QRL.2 $\mathcal{F}_i$ is a finite set consisting of differentiable *increasing convex* functions defined on an open set containing $[0, q_{\max}]$ such that $\min_{\{f_i \in \mathcal{F}_i\}} f_i(0) > 0$ and $\max_{\{f_i \in \mathcal{F}_i\}} (f_i)'(q_{\max})$ is finite.

QRL.3 $\mathcal{L}_i$ is a finite set of positive real numbers.

As indicated in Assumption QRL.1, we model the evolution of QR tradeoffs and segment lengths of each video client $i \in \mathcal{N}$ as a stationary ergodic process. Let $\left(\pi_i^{\mathcal{F},\mathcal{L}}(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}$ denote the associated marginal distribution. Without loss of generality, we assume that $\pi_i^{\mathcal{F},\mathcal{L}}(f_i, l_i) > 0$ for each $(f_i, l_i) \in \mathcal{FL}_i$. Next, let $f_{\min} := \min_{\{i, \in \mathcal{N}, f_i \in \mathcal{F}_i\}} f_i(0)$ which is strictly positive from QRL.2, and this gives a lower bound on segment compression rates. Even at zero quality, there is usually overhead information associated with a representation of a segment which causes $f_{\min}$ to be positive. The constant $q_{\max}$ represents the maximum quality that can achieved in the given network setting. Let $f_{\max} := \max_{\{i, \in \mathcal{N}, f_i \in \mathcal{F}_i\}} f_i(q_{\max})$ denote an upper bound on segment compression rates. From assumption QRL.3, it follows that $l_{\min} := \min_{\{i \in \mathcal{N}, l_i \in \mathcal{L}_i\}} l_i$ is strictly positive, and $l_{\max} := \max_{\{i \in \mathcal{N}, l_i \in \mathcal{L}_i\}} l_i$ is finite although it can be arbitrarily large.

Each video client downloads the segments of its video sequentially, and we index the segments using variables like s, s_i etc taking values in $\{0, 1, 2, ...\}$. Let $q_{i,s}$ denote the quality (i.e., STQ) associated with the segment s downloaded by video client i.

Next, we discuss our model for QoE. Our QoE model requires that the rebuffering constraint referred to in (3.2) (discussed in more detail below) are met and under this condition, QoE of video client $i \in \mathcal{N}$ depends only on the quality $(q_i)_{1:S}$ seen over segments by the video client $i \in \mathcal{N}$. Thus, our QoE model maps quality seen by a video client $i \in \mathcal{N}$ over S segments, i.e. $(q_i)_{1:S}$, to a QoE metric (under the assumption that the rebuffering constraint of video client i is met). While accurate QoE models are typically very complex, we use a simple model motivated by the discussion in Section 3.1 and the model proposed in [59]. Let $m_i^S(q_i)$ and $\text{Var}_i^S(q_i)$ denote (length weighted) mean quality and temporal variance in quality respectively associated with the first S segments downloaded by the video client i, i.e.,

$$m_i^S(q_i) := \frac{\sum_{s=1}^S l_{i,s} q_{i,s}}{\sum_{s=1}^S l_{i,s}}, \tag{3.4}$$

$$\text{Var}_i^S(q_i) := \frac{\sum_{s=1}^S l_{i,s}\left(q_{i,s} - m_i^S(q_i)\right)^2}{\sum_{s=1}^S l_{i,s}}. \tag{3.5}$$

Note that the arguments of m_i^S and Var_i^S are actually $S-$length sequences $(q_i)_{1:S}$ (i.e., $(q_{i,s})_{1\leq s\leq S}$) although we are using a shorthand for simplicity. We model the QoE of video client i for these S segments as

$$e_i^S(q_i) = m_i^S(q_i) - U_i^V\left(\text{Var}^S(q_i)\right), \tag{3.6}$$

where $\left(U_i^V(.)\right)_{i\in\mathcal{N}}$ are 'nice' convex functions satisfying assumption U.V given below:

Assumption U.V: (Variability penalty)

U.V For each $i \in \mathcal{N}$, U_i^V is a continuously differentiable increasing convex function with Lipschitz continuous derivatives defined on an open set containing $[0, q_{\max}^2]$ satisfying $\left(U_i^V\right)'(0) > 0$.

Thus, we could choose $U_i^V(v) = \eta_i v$ or $U_i^V(v) = \eta_i v^2$, where $\eta_i > 0$ and scales the penalty for temporal variability in quality. Note that our approach can be extended to more general QoE models, and we discuss this in Section 3.6.1.

Now that we have developed our QoE model, we express the objective function (3.1) in OPT-BASIC as

$$\phi_S\left((\mathbf{q})_{1:S}\right) := \sum_{i\in\mathcal{N}} U_i^E\left(e_i^S(q_i)\right) \tag{3.7}$$

where $e_i^S(q_i)$ is defined in (3.6), $\left(U_i^E(.)\right)_{i\in\mathcal{N}}$ are 'nice' concave functions satisfying assumption U.E described below. Let $e_{\min,i}:=-U_i^V(q_{\max}^2)$ and $e_{\max,i}:=q_{\max}-U_i^V(0)$.

Assumption U.E: (Fairness in QoE)

U.E For each $i \in \mathcal{N}$, we assume that U_i^E is a continuously differentiable increasing concave function with Lipschitz continuous derivatives defined on an open set containing $[e_{\min,i}, e_{\max,i}]$ satisfying $\left(U_i^E\right)'(e_{\max,i}) > 0$.

For each $i \in \mathcal{N}$, although U_i^E has to be defined over an open set containing $[e_{\min,i}, e_{\max,i}]$, only the definition of the function over $\left[-U_i^V(0), e_{\max,i}\right]$ affects the optimization. This is because we can achieve this value of QoE for each video client by just picking representation corresponding to zero quality for each segment. Thus, for example, we can choose any function from the following class of strictly concave increasing functions parametrized by $\alpha \in (0,\infty)$ ([34])

$$U_\alpha(e) = \begin{cases} \log(e) & \text{if } \alpha = 1, \\ (1-\alpha)^{-1} e^{1-\alpha} & \text{otherwise,} \end{cases} \tag{3.8}$$

and can satisfy the above conditions by making minor modifications to the function. For instance, we can use the following modification $U^{E,log}$ of the log function for any (small) $\delta > 0$: $U^{E,log}(e) = \log(e - e_{\min,i} + \delta)$, $e \in [e_{\min,i}, e_{\max,i}]$. The above class of functions are commonly used to enforce fairness specifically to achieve allocations that are $\alpha-$fair (see [43]). A larger α corresponds to a more fair allocation which eventually converges to max-min fair allocation as α goes to infinity.

Next, we consider the rebuffering related constraint considered in (3.2) of OPT-BASIC. Let $\kappa > 0$ and let $K_S = \lceil \kappa S \rceil$. We obtain a good estimate for the fraction of time spent rebuffering by a video client under an additional assumption on resource allocation that for each video client i, $\frac{1}{K_S}\sum_{k=1}^{K_S} r_{i,k}$ converges (for almost all sample paths), and hence provides an asymptotically accurate estimate for time-average resource allocation to video client i as S goes to infinity. Note that this condition is satisfied by alpha-fair resource allocation policies like proportionally fair allocation, max-min fair allocation etc. Next, note that the cumulative size of the first S segments is given by $\sum_{s=1}^{S} l_{i,s} f_{i,s}\left(q_{i,s}\right)$. Thus, a good estimate (for large S) for the time required by video client i to download the first S segments is

$$\frac{\sum_{s=1}^{S} l_{i,s} f_{i,s}\left(q_{i,s}\right)}{\frac{1}{\tau_{slot} K_S}\sum_{k=1}^{K_S} r_{i,k}}$$

which is the ratio of the cumulative size of S segments to the per slot allocation estimate. In the above observation, we are implicitly assuming that the network always has video data to send to the video client. Now, we show that the following expression is an asymptotically (as S goes to infinity) accurate estimate for the percentage of time that video client i is rebuffering while watching the S segments:

$$\beta_{i,S}\left(\left(q_i\right)_{1:S}, \left(r_i\right)_{1:K_S}\right) := \frac{\frac{\sum_{s=1}^{S} l_{i,s} f_{i,s}\left(q_{i,s}\right)}{\frac{1}{\tau_{slot} K_S}\sum_{k=1}^{K_S} r_{i,k}}}{\sum_{s=1}^{S} l_{i,s}} - 1.$$

Note that the first term in the right hand side is the ratio of the estimate for time required for download of the first S segments to the total duration $\sum_{s=1}^{S} l_{i,s}$ associated with the S segments. For video client i, let $T_i^{reb}(t)$ denote the fraction of time spent rebuffering till time $t \geq 0$ (measured in seconds), and let $T_i^{dow}(S)$ denote the time required to download S segments. Then, we have

$$T_i^{reb}(t) = \int_0^t I\left(T_i^{dow}(S_i^{seg}(u)) > u + T_i^{reb}(u)\right) du \tag{3.9}$$

where $I(.)$ is the indicator function, and $S_i^{seg}(t) = \min\left\{S : \sum_{s=1}^{S} l_{i,s} \geq t\right\}$ denotes the number of segments corresponding to video duration of t. Rearranging (3.9), we have

$$T_i^{reb}(t) = \int_0^t I\left(I_i^{reb}(u) > 0\right) du \tag{3.10}$$

where

$$I_i^{reb}(t) = \left(\frac{T_i^{dow}(S_i^{seg}(t))}{t} - 1\right) - \frac{T_i^{reb}(t)}{t}. \tag{3.11}$$

Using (3.10) and (3.11), we can show that when $I_i^{reb}(t) < 0$, ($T_i^{reb}(t)$ is non-increasing and hence) $I_i^{reb}(t)$ is non-decreasing and strictly increasing over a large enough window of time (of duration greater than $l_{\max}$) due to presence of the term $T_i^{dow}(S_i^{seg}(t)$ in (3.11). Using this observation along with the fact that $I_i^{reb}(t) \leq 0$ (since $T_i^{reb}(t) \geq T_i^{dow}(S_i^{seg}(t)) - t$ for any t), we can conclude that

$$\lim_{t\to\infty} I_i^{reb}(t) = 0.$$

Using the above observation (and set $t = \sum_{s=1}^{S} l_{i,s}$ in (3.11)) along with the convergence of $\frac{1}{K_S}\sum_{k=1}^{K_S} r_{i,k}$, we can show that that $\beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right)$ is an asymptotically accurate estimate for the percentage of time that video client i is rebuffering while watching S segments.

Note that $\beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right)$ can also take negative values which happens when segments are being downloaded at rate higher than the rate at which they are viewed. We express the rebuffering constraint in OPT-BASIC as

$$\beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right) \leq \overline{\beta}_i, \ \forall\, i \in \mathcal{N},$$

where each video client i specifies an upper bound $\overline{\beta}_i > -1$ on the percentage of time spent rebuffering. Though setting $\overline{\beta}_i = 0$ ensures that there is only an asymptotically negligible amount of rebuffering, we can enforce more stringent constraints on rebuffering by setting $\overline{\beta}_i$ to negative values.

Next, we consider the cost constraint considered in (3.2) of OPT-BASIC. The average compression rate associated with the first S segments of video client $i \in \mathcal{N}$ is $\frac{\sum_{s=1}^{S} l_{i,s} f_{i,s}(q_{i,s})}{\sum_{s=1}^{S} l_{i,s}}$. Let p_i^d denote the cost per unit of data (measured in dollar per bit) that video client $i \in \mathcal{N}$ (or the video content provider associated with the video client) has to pay. Then, the average amount of money per unit video duration the video client (/content provider) pays is

$$p_{i,S}\left((q_i)_{1:S}\right) := p_i^d \frac{\sum_{s=1}^{S} l_{i,s} f_{i,s}\left(q_{i,s}\right)}{\sum_{s=1}^{S} l_{i,s}}.$$

We express the cost constraint in OPT-BASIC as

$$p_{i,S}\left((q_i)_{1:S}\right) \leq \overline{p}_i, \;\forall\; i \in \mathcal{N},$$

where each video client i (or the video content provider associated with the video client) sets an upper bound $\overline{p}_i > 0$ on the amount of money per unit video duration.

Rest of the chapter is devoted to the derivation and analysis of an algorithm for solving OPT-BASIC that carries out *jointly* optimal quality adaptation (i.e., picks optimal $(q_i)_{1:S}$ for each video client $i \in \mathcal{N}$) and resource allocation (i.e., picks optimal $(\mathbf{r})_{1:K_S}$).

3.3 Offline optimization formulation

In this section, we formulate the problem OPT-BASIC of joint optimization of quality adaptation and resource allocation as an offline optimization problem. In the offline setting we assume $(c_k)_k$ and $(l_{i,s}, f_{i,s})_s$, i.e., the realization of the processes $(C_k)_k$ and $(L_{i,s}, F_{i,s})_s$, for each video client $i \in \mathcal{N}$ are known.

Based on the discussion in Section 3.2, we rewrite OPT-BASIC as the optimization problem OPT(S) given below:

$$
\begin{aligned}
\max_{(\mathbf{q})_{1:S},(\mathbf{r})_{1:K_S}} \quad & \phi_S\left((\mathbf{q})_{1:S}\right) \\
\text{subject to} \quad & 0 \le q_{i,s} \le q_{\max}\ \forall\ s \in \{1,...,S\}, \forall\ i \in \mathcal{N}, \\
& r_{i,k} \ge r_{i,\min},\ \forall\ k \in \{1,...,K_S\}, \forall\ i \in \mathcal{N}, \\
& c_k\left(\mathbf{r}_k\right) \le 0, \quad \forall\ k \in \{1,...,K_S\}, \\
& \beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right) \le \overline{\beta}_i, \forall\ i \in \mathcal{N}, \\
& p_{i,S}\left((q_i)_{1:S}\right) \le \overline{p}_i, \forall\ i \in \mathcal{N}.
\end{aligned}
\tag{3.12}
$$

Although the objective function of OPT(S) does not depend directly on the allocated resources $(\mathbf{r})_{1:K_S}$, the constraint (3.12) ties the quality adaptation of video clients (and hence the objective function) to their respective network resource allocation since the constraint (3.12) for video client $i \in \mathcal{N}$ is equivalent to

$$\frac{1}{\left(1+\overline{\beta}_i\right)} \frac{\sum_{s=1}^{S} l_{i,s} f_{i,s}\left(q_{i,s}\right)}{\sum_{s=1}^{S} l_{i,s}} \le \frac{1}{\tau_{slot} K_S} \sum_{k=1}^{K_S} r_{i,k}.$$

We need the following assumption to ensure strict feasibility which will be used in later sections.

Assumption-SF (Strict Feasibility): For each $c \in \mathcal{C}$, $c\left((r_{i,\min})_{i\in\mathcal{N}}\right) < 0$, and for each $i \in \mathcal{N}$, $\max_{\{f_i \in \mathcal{F}_i\}} \frac{\tau_{slot} f_i(0)}{r_{i,min}} < 1$, and $p_i^d \max_{\{f_i \in \mathcal{F}_i\}} f_i\left(0\right) < \overline{p}_i$.

This assumption[1] requires that the resource allocation $(r_{i,\min})_{i\in\mathcal{N}}$ is strictly feasible for any $c \in \mathcal{C}$, and that the maximum size of segments at zero quality is not too large.

Let ϕ_S^{opt} denote the optimal value of objective function of OPT(S). We would solve the optimization problem OPT(S) directly if it were possible. However this

[1]The assumption requires a uniform upper bound on the size of the segments at zero quality which is used in Lemma 3.1. We conjecture that this per segment requirement can be replaced with a milder averaged version.

is impossible in practice since we need to know $(c_k)_k$ and $(f_{i,s})_s$ ahead of time. Further, a direct approach is also computationally prohibitive as the optimization is over $O(NS)$ variables. Thus, from a practical point of view, the main challenge is to overcome these two hurdles and obtain a *simple* and *online* algorithm that performs as well as ϕ_S^{opt} asymptotically. We present our solution to this challenge in the next section.

3.4 An online algorithm for jointly optimizing resource allocation and quality adaptation

In this section, we present our algorithm Network Optimization for Video Adaptation (NOVA), and discuss its asymptotic optimality. The algorithm NOVA comprises three components:

1. *Allocate:* Network resource allocation is done at the beginning of each slot k by solving an optimization problem RNOVA$(\mathbf{b}_k, c_k)$ which depends on the parameter $\mathbf{b}_k$ (described later in the section) and current allocation constraint c_k.

2. *Adapt:* When a video client $i \in \mathcal{N}$ finishes download of s_ith segment, select the quality/representation for the next segment by solving an optimization problem QNOVA$_i(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$ which depends on a parameter $\boldsymbol{\theta}_{i,s_i}$ (described later in the section), and the QR tradeoff f_{i,s_i+1} of the next segment.

3. *Learn:* Learning parameters $(m_{i,s_i}, \mu_{i,s_i}, v_{i,s_i}, b_{i,k}, d_{i,s_i}, \lambda_{i,s_i})_{i\in\mathcal{N}}$ used in the optimization problems RNOVA$(\mathbf{b}_k, c_k)$ and QNOVA$_i(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$. Here s_i is the current segment index of video client i and k is current slot index. The parameters m_{i,s_i} and μ_{i,s_i} track mean quality, v_{i,s_i} tracks variance in quality, and λ_{i,s_i} tracks the mean segment duration of video client $i \in \mathcal{N}$. The

parameters $b_{i,k}$ and d_{i,s_i} serve as indicators of risk of violation of rebuffering constraints (3.12) and cost constraints (3.13) respectively of video client $i \in \mathcal{N}$, and larger the parameter, larger the risk.

We start by describing the two optimization problems RNOVA$(\mathbf{b}, c)$ and QNOVA$_i(\boldsymbol{\theta}_i, f_i)$ associated with NOVA before presenting the algorithm. We can control the response of NOVA to the indicators $b_{i,k}$ and d_{i,s_i} using functions discussed next. For each $i \in \mathcal{N}$, let $h_i^B(.)$ and $h_i^D(.)$ be non-negative valued Lipschitz continuous functions that are strictly increasing over $\mathbb{R}_+$, and are such that $\lim_{b\to\infty} h_i^B(b) = \infty$ and $\lim_{d\to\infty} h_i^D(d) = \infty$. Also, let $h_i^B(b_i) = 0$ for all $b_i \leq \underline{b}$ and $h_i^D(d_i) = 0$ for all $d_i \leq \underline{d}$ for some constants $\underline{b}$ and $\underline{d}$ typically set as zero or small negative numbers. Simple examples of functions satisfying these conditions are $\max(b, 0)$, $\max(b^2, 0)$ etc.

Let $\mathbf{b} \in \mathbb{R}^N$ and $c \in \mathcal{C}$. The optimization problem RNOVA$(\mathbf{b}, c)$ associated with network resource allocation is given below:

$$\max_{\mathbf{r}} \quad \phi^R(\mathbf{r}, \mathbf{b}) := \sum_{i \in \mathcal{N}} h_i^B(b_i)\, r_i \tag{3.13}$$

$$\text{subject to} \quad c(\mathbf{r}) \leq 0, \tag{3.14}$$

$$r_i \geq r_{i,\min} \ \forall\, i \in \mathcal{N}. \tag{3.15}$$

Let $\mathcal{R}^*(\mathbf{b}, c)$ denote the set of optimal solutions to RNOVA$(\mathbf{b}, c)$. Note that the objective function (3.13) gives more weight to video clients higher value of b_i, i.e., higher risk of violation of rebuffering constraints.

The optimization problem QNOVA$_i(\boldsymbol{\theta}_i, f_i)$ associated with quality adaptation performed by video clients. Let $0 \leq m_i, \mu_i \leq q_{\max},\ 0 \leq v_i \leq q_{\max}^2,\ b_i, d_i \in \mathbb{R}$, $\boldsymbol{\theta}_i = (m_i, \mu_i, v_i, b_i, d_i)$ and $f_i \in \mathcal{F}_i$. For $i \in \mathcal{N}$, the optimization problem

$\text{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$ is given below:

$$\begin{aligned} \max_{q_i} \quad & \phi^Q(q_i, \boldsymbol{\theta}_i, f_i) \\ \text{subject to} \quad & q_i \geq 0, \qquad (3.16) \\ & q_i \leq q_{\max}, \qquad (3.17) \end{aligned}$$

where

$$\begin{aligned} \phi^Q(q_i, \boldsymbol{\theta}_i, f_i) &= \left(U_i^E\right)' \left(\mu_i - U_i^V(v_i)\right)\left(q_i - \left(U_i^V\right)'(v_i)(q_i - m_i)^2\right) \qquad (3.18) \\ &\quad - \frac{h_i^B(b_i)}{\left(1+\overline{\beta}_i\right)} f_i(q_i) - \frac{p_i^d h_i^D(d_i)}{\overline{p}_i} f_i(q_i). \end{aligned}$$

We can obtain an intuitive understanding of the objective function (3.18) of the above optimization problem by noting that the term $(q_i - m_i)^2$ ensures that an optimal solution to $\text{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$ is not too far away from m_i (current estimate of mean quality), and thus avoids high variance in quality. Also, the terms $\frac{h_i^B(b_i)}{(1+\overline{\beta}_i)} f_i(q_i)$ and $\frac{p_i^d h_i^D(d_i)}{\overline{p}_i} f_i(q_i)$ in (3.18) penalize quality choices leading to large segment sizes when b_i or d_i are high, and thus help NOVA to respond to increased risk of violation of rebuffering constraints and cost constraints. The optimization problem $\text{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$ is convex and has a unique solution due to the strict concavity of the objective function. Let $q_i^*(\boldsymbol{\theta}_i, f_i)$ denote the solution to $\text{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$.

For each $i \in \mathcal{N}$, define the set $\mathcal{H}^{(i)}$ as follows:

$$\begin{aligned} \mathcal{H}^{(i)} &= \{(m_i, \mu_i, v_i, b_i, d_i, \lambda_i) : \ 0 \leq m_i, \mu_i \leq q_{\max}, \ 0 \leq v_i \leq q_{\max}^2, \qquad (3.19) \\ &\quad b_i \geq \underline{b}, \ d_i \geq \underline{d}, \ l_{\min} \leq \lambda_i \leq l_{\max}\}. \end{aligned}$$

Let s_i be an indexing variable keeping track of the segment video client i is currently downloading. We also use auxiliary variables $b_{Q,i,s}$ and $b_{R,i,k}$ for each video client i to keep track of the parameter $b_{i,\cdot}$ in NOVA. The algorithm NOVA is given below.

NOVA

NOVA.0: Initialization: Let $(m_{i,0}, \mu_{i,0}, v_{i,0}, b_{i,0}, d_{i,0}, \lambda_{i,0}) \in \mathcal{H}^{(i)}$ and $s_i = 0$ for each $i \in \mathcal{N}$, and $\epsilon > 0$.

In each slot $k \geq 0$, carry out the following steps:

ALLOCATE: At the beginning of slot k, let $b_{R,i,k} = b_{i,k}$ for each $i \in \mathcal{N}$, and allocate resources according to any element of the set $\mathcal{R}^* (\mathbf{b}_k, c_k)$ (of optimal solutions to RNOVA($\mathbf{b}_k, c_k$)) and update $\mathbf{b}_k$ as follows:

$$b_{i,k+1} = b_{i,k} + \epsilon \left(\frac{\tau_{slot}}{(1 + \overline{\beta}_i)} \right). \tag{3.20}$$

ADAPT: In slot k, if any video client $i \in \mathcal{N}$ finishes download of s_i th segment, let $b_{Q,i,s_i+1} = b_{i,k+1}$, $\boldsymbol{\theta}_{i,s_i} = (m_{i,s_i}, \mu_{i,s_i}, v_{i,s_i}, b_{Q,i,s_i+1}, d_{i,s_i})$. For segment $s_i + 1$ of video client i, select representation with quality $q_i^*(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$ (i.e., optimal solution to QNOVA$_i(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$), denoted as q^*_{i,s_i+1} for brevity, and update parameters m_{i,s_i+1}, μ_{i,s_i+1}, v_{i,s_i+1}, $b_{i,k+1}$, d_{i,s_i+1} and s_i as follows:

$$m_{i,s_i+1} = m_{i,s_i} + \epsilon \left(U_i^E\right)' \left(\mu_i - U_i^V(v_i)\right) \left(U_i^V\right)' (v_i) \left(\frac{l_{i,s_i+1}}{\lambda_{i,s_i}} q^*_{i,s_i+1} - m_{i,s_i} \right) \tag{3.21}$$

$$\mu_{i,s_i+1} = \mu_{i,s_i} + \epsilon \left(\frac{l_{i,s_i+1}}{\lambda_{i,s_i}} q^*_{i,s_i+1} - \mu_{i,s_i} \right), \tag{3.22}$$

$$v_{i,s_i+1} = v_{i,s_i} + \epsilon \left(\frac{l_{i,s_i+1}}{\lambda_{i,s_i}} \left(q^*_{i,s_i+1} - m_{i,s_i}\right)^2 - v_{i,s_i} \right), \tag{3.23}$$

$$b_{i,k+1} = \left[b_{i,k+1} - \epsilon \left(l_{i,s_i+1}\right)\right]_{\underline{b}}, \tag{3.24}$$

$$d_{i,s_i+1} = \left[d_{i,s_i} + \epsilon \left(p_i^d \frac{l_{i,s_i+1} f_{i,s_i+1}\left(q^*_{i,s_i+1}\right)}{\overline{p}_i} - \lambda_{i,s_i} \right) \right]_{\underline{d}}, \tag{3.25}$$

$$\lambda_{i,s_i+1} = \lambda_{i,s_i} + \epsilon \left(l_{i,s_i+1} - \lambda_{i,s_i}\right), \tag{3.26}$$

$$s_i = s_i + 1.$$

Here, $[x]_y = \max(x, y)$ for $x, y \in \mathbb{R}$. The variable $b_{Q,i,s}$ stores the value of $b_{i,\cdot}$ used in choosing quality for the sth segment of video client i, and $b_{R,i,k}$ stores the value of $(b_{i,k})_{i \in \mathcal{N}}$ used in the resource allocation in slot k. These are just auxiliary variables, and do not affect the evolution of the algorithm (unlike $b_{i,k}$ which affects the algorithm). Further, to ensure that the video clients start downloading video segments from the beginning, we assume that all the video clients have already downloaded 0th segment.

Allocation and adaptation in NOVA are *asynchronous*- allocation is done at the beginning of each slot, and adaptation related decisions are made when video clients complete a segment download. The update equation (3.26) associated with the parameter λ_{i,s_i} is similar to update rules used for tracking EWMA (Exponentially Weighted Moving Averages), and ensures that λ_{i,s_i} tracks the mean segment duration of video client i. The update rules (3.21)-(3.23) are similar, and ensure that m_{i,s_i} and μ_{i,s_i} track mean quality, while v_{i,s_i} tracks variance in quality. Both m_{i,s_i} and μ_{i,s_i} track mean quality giving different weights to the current quality, and we later generalize the update rule (3.22) so that μ_{i,s_i} tracks parameters associated with more general QoE metrics. The weights $\frac{l_{i,s_i+1}}{\lambda_{i,s_i}}$ used in the update rules ensure that the duration of the segment is appropriately factored. Next, we consider the evolution of the parameter the operator $b_{i,k}$ which is updated in both (3.20) and (3.24) ignoring $[.]_{\underline{b}}$ and setting initialization to zero. (3.20) ensures that $b_{i,k}$ is increased by fixed amount $\epsilon \left(\frac{\tau_{slot}}{(1+\overline{\beta}_i)} \right)$ at the beginning of each slot. (3.24) ensures that when a video client completes the download of a segment, $b_{i,k}$ is reduced by ϵ times the duration of the next segment. Hence, at some time t seconds (or t/τ_{slot} slots) after starting the video,

$$\frac{b_{i,k}}{\epsilon} \approx \left(\frac{t}{(1+\overline{\beta}_i)} - \text{Duration of video downloaded till now} \right),$$

which sheds light on its role as an indicator of risk of violation of rebuffering constraints (3.12) for video client i, for e.g., $b_{i,k}$ will be large if $\overline{\beta}_i = 0$ and the total duration of video downloaded till now is much less than t. Similarly, we can argue that d_{i,s_i} serves as an indicator of risk of violation of cost constraint (3.13) for video client i. Depending on the problem under consideration, we can drop some of the parameters from NOVA. For instance, if U_i^V is a linear function, we need not track v_{i,s_i}. Or, if a video client does not have a cost constraint, we need not track d_{i,s_i}.

It is interesting to note that the quality adaptation proposed in NOVA does not directly use any information about the allocation constraints. Neither does the resource allocation directly use any information about QR tradeoffs of the video clients. Yet, the joint resource allocation and quality adaptation under NOVA has strong optimality properties (which are presented later in this section). This is mainly due to the fact that the variables $(b_{i,k})_{i\in\mathcal{N}}$ carry almost all the information about the video clients' quality adaptation that is required by the network controller to carry out optimal resource allocation, and the variable $b_{i,k}$ carries almost all the information that the quality adaptation at video client i needs to know about the resource allocation (to the client). For e.g., consider a video client i in the network that has very few unwatched segments in the playback buffer, i.e., the video client is about to experience rebuffering. We see that the update rules for $b_{i,k}$ (and a large enough initialization) ensure that $b_{i,k}$ will be large in this scenario, and this forces the video client and the network controller to make the right moves, i.e., this forces the video client to switch to low quality representations (while taking current QR tradeoffs into account), and forces the network controller to give higher priority to this video client in the resource allocation (while taking the current allocation constraints also into account).

Next, we discuss some important features of NOVA that make it attractive from a practical point of view.

- We provide *strong optimality guarantees* for NOVA (see Theorem 3.1).

- NOVA is an *online* algorithm as it only uses *current* information, i.e., NOVA only needs the allocation constraint c_k for slot k, and for quality adaptation of segment $s_i + 1$ of video client i, it only requires the QR tradeoff f_{i,s_i+1} for the optimization and l_{i,s_i+1} for updates associated with that segment.

- NOVA is a *simple* algorithm since RNOVA$(\mathbf{b}, c)$ is convex optimization problem in N variables. Further, if allocation constraints are linear, RNOVA$(\mathbf{b}, c)$ is just a linear program which often has enough structure to allow for very efficient evaluation of the solutions. Also, note that QNOVA$_i(\boldsymbol{\theta}_i, f_i)$ is just a scalar convex optimization problem.

- The asynchronous nature of NOVA ensures that the video clients can work at their own pace and the adaptation prescribed in NOVA is entirely *client driven* requiring no assistance from the network controller, and is thus well suited for DASH framework.

- NOVA can be implemented in a *distributed* manner with minimal signaling since quality adaptation is client driven and for the resource allocation, the network controller need only know $\mathbf{b}_k$ which are indicators of risk of violation of rebuffering constraints associated with the video clients (illustrated in Fig. 3.1). To ensure that the network controller knows the current value of $\mathbf{b}_k$, each video client can send a signal to the base station indicating the latest value of $b_{i,k}$ at the end of each segment download which usually occurs at a low frequency (typically once a second). On receiving this signal from video client $i \in \mathcal{N}$, the network controller can then update $b_{i,k}$ until the next signal from video client i using the simple update rule in (3.20) that requires only constant increments.

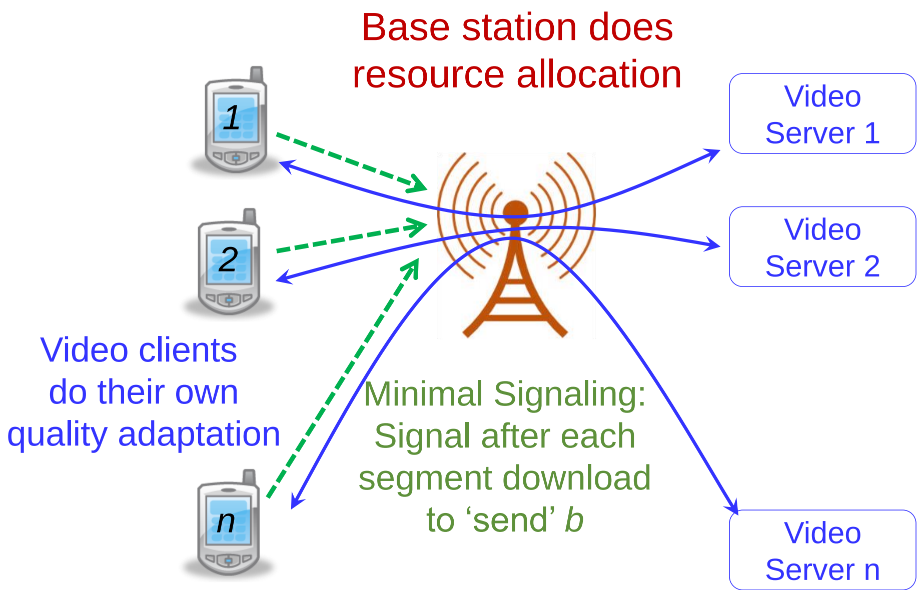

Figure 3.1: Distributed Implementation of NOVA

- Optimization algorithm for resource allocation, RNOVA$(\mathbf{b}, c)$ requires only a simple modification of legacy schedulers like proportionally fair schedulers (see [30]). In fact, the optimization problem associated with proportionally fair schedulers is almost the same as RNOVA$(\mathbf{b}, c)$ except that it uses a function of current estimate of average throughput instead of $\mathbf{b}_k$.

The preceding discussion of NOVA suggests it is intuitively doing the right things. The discussion in the rest of this section and Section 3.5 is a rigorous analysis of NOVA aimed at establishing the strong optimality result for NOVA given in Theorem 3.1. Proving Theorem 3.1 requires other intermediary results. We have devoted Section 3.5 to these results and give a proof of Theorem 3.1 at the end of that section.

Theorem 3.1. *Suppose* $(m_{i,\cdot}, \mu_{i,\cdot}, v_{i,\cdot}, b_{i,\cdot}, d_{i,\cdot}, \lambda_{i,\cdot})$ *evolve according to the update rules in NOVA.*

(a) Feasibility: *NOVA asymptotically satisfies the constraints on rebuffering and*

cost, i.e., for each $i \in \mathcal{N}$

$$\begin{aligned} limsup_{S\to\infty}\beta_{i,S}\left((q_i^*)_{1:S},(r_i^*)_{1:K_S}\right) &\leq \overline{\beta}_i, &(3.27)\\ limsup_{S\to\infty}p_{i,S}\left((q_i^*)_{1:S}\right) &\leq \overline{p}_i. &(3.28)\end{aligned}$$

(b) Optimality: *Let* $S_\epsilon = \frac{S}{\epsilon}$. *Then,*

$$\lim_{S\to\infty}\lim_{\epsilon\to 0}\left(\phi_{S_\epsilon}\left(\left((q_i^*\left(\boldsymbol{\theta}_{i,s},f_{i,s}\right))_{i\in\mathcal{N}}\right)_{1\leq s\leq S_\epsilon}\right)-\phi_{S_\epsilon}^{opt}\right)$$

goes to zero in probability.

The above result tells us that the difference in performance of the *online* algorithm NOVA (i.e., $\phi_{S_\epsilon}\left((\mathbf{q}^*)_{1:S_\epsilon}\right)$) and that of the optimal *offline* scheme goes to zero for long enough videos and small enough ϵ. Recall that $\phi_{S_\epsilon}^{opt}$ is the optimal value of the OPT(S_ϵ), i.e., the performance of the optimal omniscient offline scheme which knows all the allocation constraints $(c_k)_k$ and QR tradeoffs and segment lengths $(f_{i,s}, l_{i,s})_s$ ahead of time. Note that although choosing small ϵ is beneficial for long videos, it can significantly affect the performance (initial transient and tracking ability) of NOVA for short videos.

In the rest of this section, we discuss some useful properties of NOVA that will be used in Section 3.5. We start with optimality conditions associated with solutions to RNOVA($\mathbf{b}, c$) and QNOVA$_i(\boldsymbol{\theta}_i, f_i)$. The optimization problem RNOVA($\mathbf{b}, c$) is convex, and using Assumption-SF, we can show that it satisfies Slater's condition (see [10] for reference). Thus, KKT conditions are necessary (and sufficient) for optimality. Hence, if $\mathbf{r}^*(\mathbf{b}, c)$ is an optimal solution to RNOVA($\mathbf{b}, c$), there exist constants $\chi^*(c)$ and $(\omega_i^*(c))_{i\in\mathcal{N}}$ such that for each $i \in \mathcal{N}$,

$$h_i^B(b_i) = \chi^*(c)c_i^{'}(\mathbf{r}^*(\mathbf{b},c)) + \omega_i^*(c), \tag{3.29}$$

$$\chi^*(c)c(\mathbf{r}^*(\mathbf{b},c)) = 0, \tag{3.30}$$

$$\omega_i^*(c)(r_i^*(\mathbf{b},c) - r_{i,\min}) = 0, \tag{3.31}$$

The optimization problem $\text{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$ is also convex and satisfies Slater's condition (since the constraints are all linear), and thus, KKT conditions are necessary (and sufficient) for optimality. Thus, there exist constants $\gamma_i(\boldsymbol{\theta}_i, f_i)$ and $\overline{\gamma}_i(\boldsymbol{\theta}_i, f_i)$ such that

$$\begin{aligned} &\left(U_i^E\right)^{'}\left(\mu_i - U_i^V(v_i)\right)\left(1 - 2\left(U_i^V\right)^{'}(v_i)\left(q_i^*(\boldsymbol{\theta}_i, f_i) - m_i\right)\right) + \gamma_i(\boldsymbol{\theta}_i, f_i) \\ &-\overline{\gamma}_i(\boldsymbol{\theta}_i, f_i) - \frac{h_i^B(b_i)}{\left(1+\overline{\beta}_i\right)}(f_i)^{'}\left(q_i^*(\boldsymbol{\theta}_i, f_i)\right) - p_i^d\frac{h_i^D(d_i)}{\overline{p}_i}(f_i)^{'}\left(q_i^*(\boldsymbol{\theta}_i), f_i\right) = 0, \end{aligned} \tag{3.32}$$

$$\gamma_i(\boldsymbol{\theta}_i, f_i)\, q_i^*(\boldsymbol{\theta}_i, f_i) = 0, \tag{3.33}$$

$$\overline{\gamma}_i(\boldsymbol{\theta}_i, f_i)\left(q_i^*(\boldsymbol{\theta}_i, f_i) - q_{\max}\right) = 0. \tag{3.34}$$

The next result states that the parameters in NOVA stay in a compact set.

Lemma 3.1. *For any initialization* $(m_{i,0}, \mu_{i,0}, v_{i,0}, b_{i,0}, d_{i,0}, \lambda_{i,0})_{i\in\mathcal{N}} \in \prod_{i\in\mathcal{N}} \mathcal{H}^{(i)}$, *the parameters evolving according to NOVA satisfy the following: for each* $i \in \mathcal{N}$, $s \geq 1$ *and* $k \geq 1$, *we have* $0 \leq m_{i,s}, \mu_{i,s} \leq q_{\max}$, $0 \leq v_{i,s} \leq q_{\max}^2$, *and* $l_{\min} \leq \lambda_{i,s} \leq l_{\max}$. *Further,* $\underline{b} \leq b_{i,k} \leq \overline{b}$, $\underline{d} \leq d_{i,s} \leq \overline{d}$ *for some finite constants* $\overline{b}$ *and* $\overline{d}$ *and for all* k *and* s *large enough.*

Proof. It is easy to establish the result for the parameters $m_{i,s}, \mu_{i,s}, v_{i,s}$ and $\lambda_{i,s}$ using the initialization of these parameters in NOVA and the boundedness of the quantities involved in the respective update rules. For instance, we can use (3.21), (3.22) and the fact that $0 \leq q_{i,s_i+1}^* \leq q_{\max}$ to obtain the result for $m_{i,\cdot}$ and $\mu_{i,\cdot}$.

Next, we show that there exists a finite $\overline{b}$ such that $\underline{b} \leq b_{i,k} \leq \overline{b}$ for all k

large enough. The lower bound is easy to show and holds for all k. We establish the upper bound by showing the following property regarding the optimal solution to QNOVA$_i(\boldsymbol{\theta}_i, f_i)$: for each $i \in \mathcal{N}$, there is a finite constant $\overline{b}_i$ such that

$$\max_{\{f_i \in \mathcal{F}_i\}} \frac{f_i\left(q_i^*\left(\boldsymbol{\theta}_i, f_i\right)\right)}{\left(1+\overline{\beta}_i\right)} - \frac{r_{i,min}}{\tau_{slot}} \leq 0.5 \left(\max_{\{f_i \in \mathcal{F}_i\}} \frac{f_i\left(0\right)}{\left(1+\overline{\beta}_i\right)} - \frac{r_{i,min}}{\tau_{slot}} \right)$$

for any $\boldsymbol{\theta}_i = (m_i, \mu_i, v_i, b_i, d_i)$ satisfying $b_i \geq \overline{b}_i$, $0 \leq m_i, \mu_i \leq q_{\max}$, $0 \leq v_i \leq q_{\max}^2$, $d_i \geq \underline{d}$. Note that $\max_{\{f_i \in \mathcal{F}_i\}} \frac{f_i(0)}{\left(1+\overline{\beta}_i\right)} - \frac{r_{i,min}}{\tau_{slot}} < 0$ from Assumption-SF. Hence, if this property holds, we can conclude that for large enough $b_{i,k}$, i.e. $b_{i,k} \geq \overline{b} := \max_{i \in \mathcal{N}} \max\left(\overline{b}_i\right)$, the time required to download a segment is strictly less than $\left(1+\overline{\beta}_i\right)$ times the duration of video associated with the segment. Thus, $b_{i,k}$ is strictly decreasing when it is greater than $\overline{b}$. Hence for any initialization $b_{i,k} \geq \underline{b}$, we can show that $b_{i,k} \geq \overline{b}$ for large enough k.

We establish the above property next. Using (3.32) and the fact that f_i are convex increasing functions, we have

$$\begin{aligned} &\left(U_i^E\right)' \left(\mu_i - U_i^V\left(v_i\right)\right) \left(1 - 2\left(U_i^V\right)'\left(v_i\right) q_i^*\left(\boldsymbol{\theta}_i, f_i\right)\right) \\ &\qquad\qquad \geq \frac{h_i^B\left(b_i\right)}{\left(1+\overline{\beta}_i\right)} \left(f_i\right)' \left(q_i^*\left(\boldsymbol{\theta}_i, f_i\right)\right) - \gamma_i\left(\boldsymbol{\theta}_i, f_i\right). \end{aligned}$$

Let

$$\begin{aligned} \eta_1 &= \max_{i \in \mathcal{N}} \max_{e_i \in [e_{\min,i}, e_{\max,i}]} \left(1+\overline{\beta}_i\right) \left(U_i^E\right)'\left(e_i\right), \\ \eta_2 &= \min_{i \in \mathcal{N}} \min_{e_i \in [e_{\min,i}, e_{\max,i}], v_i \in [0, q_{\max}^2]} 2\left(1+\overline{\beta}_i\right) \left(U_i^E\right)'\left(e_i\right) \left(U_i^V\right)'\left(v_i\right). \end{aligned}$$

Recall that $e_{\min,i} = -U_i^V(q_{\max}^2)$ and $e_{\max,i} = q_{\max} - U_i^V(0)$. Since, $\left(U_i^E\right)'(.)$ and $\left(U_i^V\right)'(.)$ are continuous, $\left(U_i^V\right)'(0) > 0$ and $\left(U_i^E\right)'(e_{\max,i}) > 0$, η_1 is finite, and

$\eta_2 > 0$. Hence, for any $i \in \mathcal{N}$ and $\boldsymbol{\theta}_i$

$$\eta_1 - \eta_2 q_i^* (\boldsymbol{\theta}_i, f_i) \geq h_i^B (b_i) (f_i)' (q_i^* (\boldsymbol{\theta}_i, f_i)) - \gamma_i (\boldsymbol{\theta}_i, f_i),$$

Using the above inequality, and using the facts that $(f_i)'(q) > 0$ for each $q > 0$ and $\lim_{b \to \infty} h_i^B(b) = \infty$, we can show that $\lim_{b \to \infty} q_i^* (\boldsymbol{\theta}_i, f_i) = 0$. Also, from Assumption-SF, $\max_{\{f_i \in \mathcal{F}_i\}} \frac{f_i(0)}{(1+\overline{\beta}_i)} - \frac{r_{i,min}}{\tau_{slot}} < 0$. Now, (using continuity of the functions in $\mathcal{F}_i$ and finiteness of $|\mathcal{F}_i|$) we can conclude that there is some finite constant $\overline{b}_i$ such that $\max_{\{f_i \in \mathcal{F}_i\}} \frac{f_i(q_i^*(\boldsymbol{\theta}_i, f_i))}{(1+\overline{\beta}_i)} - \frac{r_{i,min}}{\tau_{slot}} \leq 0.5 \left(\max_{\{f_i \in \mathcal{F}_i\}} \frac{f_i(0)}{(1+\overline{\beta}_i)} - \frac{r_{i,min}}{\tau_{slot}} \right)$ when $b_i \geq \overline{b}_i$.

The proof for $d_{i,s}$ can be completed using an approach similar to that given for $b_{i,k}$. □

For the next two results, let $\boldsymbol{\theta}_i = (m_i, \mu_i, v_i, b_i, d_i)$ where $0 \leq m_i, \mu_i \leq q_{\max}$, $0 \leq v_i \leq q_{\max}^2$ and $b_i, d_i \in \mathbb{R}$. The next result provides smoothness properties for the optimal solutions of RNOVA$(\mathbf{b}, c)$ and QNOVA$_i(\boldsymbol{\theta}_i, f_i)$.

Lemma 3.2. *(a) For each $i \in \mathcal{N}$ and $f_i \in \mathcal{F}_i$, $q_i^* (\boldsymbol{\theta}_i, f_i)$ is a continuous function of $\boldsymbol{\theta}_i$.*
(b) For each $c \in \mathcal{C}$, $\mathcal{R}^* (\mathbf{b}, c)$ is a convex and compact set. Further, $\mathcal{R}^* (\mathbf{b}, c)$ is an upper semi-continuous set valued map of $\mathbf{b}$.
(c) For each $c \in \mathcal{C}$ and $\mathbf{r}^* (\mathbf{b}, c) \in \mathcal{R}^* (\mathbf{b}, c)$, $\phi^R (\mathbf{r}^* (\mathbf{b}, c), \mathbf{b})$ is a continuous function of $\mathbf{b}$.

Proof. Part (a) follows from Theorem 2.2 in [16] which provides sufficient conditions for verifying continuity of the optimal solution $q_i^* (\boldsymbol{\theta}_i, f_i)$. We can verify that the conditions given in Theorem 2.2 are satisfied since $\phi^Q (q_i, \boldsymbol{\theta}_i, f_i)$ is continuous in $(q_i, \boldsymbol{\theta}_i)$, QNOVA$_i(\boldsymbol{\theta}_i, f_i)$ has a unique solution, and since the set of feasible solutions is a compact set.

Part (b) follows from Theorem 2.4 in [16] which provides sufficient conditions for verifying continuity of the set of optimal solutions $\mathcal{R}^*(\mathbf{b}, c)$ of RNOVA($\mathbf{b}, c$)..

Part (c) follows from Theorem 2.1 in [16] which provides sufficient conditions for verifying continuity of the optimal value $\phi^R(\mathbf{r}^*(\mathbf{b}, c), \mathbf{b})$ of RNOVA($\mathbf{b}, c$). □

In the next result, we discuss concavity and differentiability properties of the optimal value of QNOVA$_i(\boldsymbol{\theta}_i, f_i)$, i.e., $\phi^Q(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i)$.

Lemma 3.3. *The following statements hold for each $i \in \mathcal{N}$ and $f_i \in \mathcal{F}_i$.*
(a) The optimal value of QNOVA$_i(\boldsymbol{\theta}_i, f_i)$, i.e., $\phi^Q(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i)$, is a strictly concave function of m_i where $\boldsymbol{\theta}_i = (m_i, \mu_i, v_i, b_i, d_i)$.
(b) The partial derivative of $\phi^Q(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i)$ with respect of m_i is given by:

$$\frac{\partial \phi^Q(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i)}{\partial m_i} = 2\left(U_i^E\right)'\left(\mu_i - U_i^V(v_i)\right)\left(U_i^V\right)'(v_i)\left(q_i^*(\boldsymbol{\theta}_i, f_i) - m_i\right). \quad (3.35)$$

(c) Let $\boldsymbol{\theta}_i^{(m)} = (m, \mu_i, v_i, b_i, d_i)$, i.e., $\boldsymbol{\theta}_i$ with the first component set to m. If $m \neq m_i$, the optimal value of QNOVA$_i(\boldsymbol{\theta}_i^{(m)}, f_i)$ satisfies

$$\begin{aligned} \phi^Q\left(q_i^*\left(\boldsymbol{\theta}_i^{(m)}, f_i\right), \boldsymbol{\theta}_i^{(m)}, f_i\right) &< \phi^Q(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i) \\ &+2(m - m_i)\left(U_i^E\right)'\left(\mu_i - U_i^V(v_i)\right)\left(U_i^V\right)'(v_i)\left(q_i^*(\boldsymbol{\theta}_i, f_i) - m_i\right). \end{aligned} \quad (3.36)$$

Proof. Part (a) follows from Proposition 2.8 from [17] which provides sufficient conditions for verifying strict concavity of the optimal value of an optimization problem with respect to parameters associated with the problem.

Part (b) follows from Theorem 4.1 related to sensitivity analysis of optimal value function given in [7], and the remark following the theorem. Part (c) follows from strict concavity in part (a) and using the expression for the partial derivative in part (b). □

3.5 Proof of optimality of NOVA

This section is devoted to the proof of the previously stated Theorem 3.1 related to optimality of NOVA. In Subsection 3.5.1, we study an auxiliary optimization problem OPTSTAT and obtain Theorem 3.2 which suggests that we can prove the main optimality result Theorem 3.1 for NOVA if we establish an appropriate convergence result for NOVA's parameters. In Subsection 3.5.2, we study an auxiliary differential inclusion (given in (3.75)-(3.82)) which evolves according to average dynamics of NOVA, and obtain a convergence result for the differential inclusion. In Subsection 3.5.3, we view NOVA's update equations ((3.21)-(3.26) and (3.72)-(3.73)) as an asynchronous stochastic approximation update (see, e.g., [29] for reference), and relate this stochastic approximation update to the auxiliary differential inclusion (in (3.75)-(3.82)), and use this relationship to establish desired convergence of NOVA's parameters using the convergence result for the auxiliary differential inclusion established in Subsection 3.5.2.

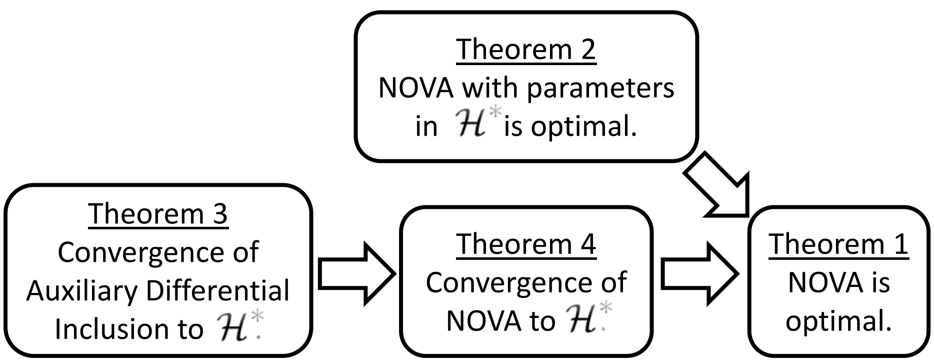

Figure 3.2: An outline of proof of optimality of NOVA

3.5.1 OPTSTAT: An auxiliary optimization problem related to the offline optimization formulation

The offline optimization formulation OPT(S) mainly involves time and segment averages of various quantities. By contrast, the formulation of OPTSTAT discussed in this section is based on the expected value of the corresponding quantities evaluated under the stationary distribution of $(C_k)_k$ and $(F_{i,s}, L_{i,s})_{s\geq 0}$ for each $i \in \mathcal{N}$.

Recall (see Section 3.2) that $(C_k)_k$ is stationary ergodic random process with marginal distribution $\left(\pi^{\mathcal{C}}(c)\right)_{c\in\mathcal{C}}$, and let C^{π} denote a random variable with distribution $\left(\pi^{\mathcal{C}}(c) : c \in \mathcal{C}\right)$. Also, recall that for each $i \in \mathcal{N}$, $(F_{i,s}, L_{i,s})_{s\geq 0}$ is a stationary ergodic process with marginal distribution $\left(\pi_i^{\mathcal{F},\mathcal{L}}(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}$. We let (F_i^{π}, L_i^{π}) denote random variables with distribution $\left(\pi_i^{\mathcal{F},\mathcal{L}}(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}$.

Let $(\mathbf{r}(c))_{c\in\mathcal{C}}$ be a vector (of vectors) representing the reward allocation $\mathbf{r}(c)$ ($\in \mathbb{R}^N$) to the video clients for each $c \in \mathcal{C}$. Although we are abusing the notation introduced earlier where $\mathbf{r}(t)$ denoted the allocation to the video clients in slot t, one can differentiate between the functions based on the context in which they are being discussed. Similarly, we let $q_i(f, l)$ denote the quality associated with a segment of video client i with $(f, l) \in \mathcal{FL}_i$. Mimicking the definition of $\phi_S\left((\mathbf{q})_{1:S}\right)$, $m_i^S(q_i)$ and $\mathrm{Var}_i^S(q_i)$ in Section 3.3, we let

$$\begin{aligned} &\phi_{\pi}\left(\left(\left(q_i(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}\right) \\ &\quad = \sum_{i\in\mathcal{N}} U_i^E\left(\mathrm{Mean}\left(q_i(F_i^{\pi}, L_i^{\pi})\right) - U_i^V\left(\mathrm{Var}\left(q_i(F_i^{\pi}, L_i^{\pi})\right)\right)\right), \end{aligned} \tag{3.37}$$

where

$$\begin{aligned} \mathrm{Mean}\left(q_i(F_i^{\pi}, L_i^{\pi})\right) &= \frac{\mathbb{E}\left[L_i^{\pi} q_i(F_i^{\pi}, L_i^{\pi})\right]}{\mathbb{E}\left[L_i^{\pi}\right]}, \\ \mathrm{Var}\left(q_i(F_i^{\pi}, L_i^{\pi})\right) &= \frac{\mathbb{E}\left[L_i^{\pi}\left(q_i(F_i^{\pi}, L_i^{\pi}) - \mathrm{Mean}\left(q_i(F_i^{\pi}, L_i^{\pi})\right)\right)^2\right]}{\mathbb{E}\left[L_i^{\pi}\right]}. \end{aligned}$$

Now, consider the optimization problem OPTSTAT given below:

$$\max_{\left((q_i(f_i,l_i))_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},(\mathbf{r}(c))_{c\in\mathcal{C}}} \quad \phi_\pi\left(\left(\left(q_i\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}\right) \tag{3.38}$$

$$\begin{aligned}
\text{subject to } c\left(\mathbf{r}\left(c\right)\right) &\leq 0,\ \forall\ c\in\mathcal{C}, && (3.39)\\
q_i\left(f_i,l_i\right) &\geq 0,\ \forall\ \left(f_i,l_i\right)\in\mathcal{FL}_i,\ \forall\ i\in\mathcal{N}, && (3.40)\\
q_i\left(f_i,l_i\right) &\leq q_{\max},\ \forall\ \left(f_i,l_i\right)\in\mathcal{FL}_i,\ \forall\ i\in\mathcal{N}, && (3.41)\\
r_i\left(c\right) &\geq r_{i,\min},\ \forall\ c\in\mathcal{C},\ \forall\ i\in\mathcal{N}, && (3.42)\\
p_i^d\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\overline{p}_i\mathbb{E}\left[L_i^\pi\right]} &\leq 1,\ \forall\ i\in\mathcal{N}, && (3.43)\\
\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} &\leq \frac{\mathbb{E}\left[r_i\left(C^\pi\right)\right]}{\tau_{slot}},\ \forall\ i\in\mathcal{N}. && (3.44)
\end{aligned}$$

We obtained the above formulation by replacing the time and segment averages of various quantities in OPT(S) (see (3.12)-(3.12)) with the expected value of the corresponding quantities. Note that in the constraint $c\left(\mathbf{r}\left(c\right)\right)\leq 0$ given in (3.39), c appearing as argument of $\mathbf{r}(c)$ is an index (for the corresponding element in $\mathcal{C}$) whereas the other c is the associated function with argument $\mathbf{r}(c)$. Similarly, in the term $F_i^\pi\left(q_i\left(F_i^\pi,L_i^\pi\right)\right)$, the argument $\left(F_i^\pi,L_i^\pi\right)$ serves as an index whereas $F_i^\pi(.)$ is the (random) function.

For $\delta\geq 0$, let OPTSTAT$_\delta$ denote a modification of optimization problem OPTSTAT with objective (3.38), constraints (3.39)-(3.42), and the following constraints

$$\begin{aligned}
p_i^d\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\overline{p}_i\mathbb{E}\left[L_i^\pi\right]} &\leq 1+\delta,\ \forall\ i\in\mathcal{N}, && (3.45)\\
\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} &\leq \frac{\mathbb{E}\left[r_i\left(C^\pi\right)\right]}{\tau_{slot}}+\delta,\ \forall\ i\in\mathcal{N}. && (3.46)
\end{aligned}$$

Hence, OPTSTAT$_\delta$ is obtained by relaxing constraints (3.43) and (3.44) of OPTSTAT by δ. Let OPTSTATVAL and OPTSTATVAL$_\delta$ denote the optimal value of

OPTSTAT and OPTSTAT$_\delta$ respectively. Clearly, for any $\delta \geq 0$, OPTSTATVAL $\leq$ OPTSTATVAL$_\delta$ and we have equality when $\delta = 0$.

The next result presents properties related to the optimal solution of OPTSTAT, and optimal values of OPTSTAT and OPTSTAT$_\delta$. Part (a) states says that OPTSTAT is a nice convex optimization problem, and part (b) states that the optimal quality choices obtained by solving OPTSTAT are unique. Part (c) shows that for any video client $i \in \mathcal{N}$, the optimal quality choices for any two segments with the same QR tradeoff are the same irrespective of their lengths. In part (d), we establish continuity of optimal value OPTSTATVAL$_\delta$ of OPTSTAT$_\delta$ at $\delta = 0$.

Lemma 3.4. *(a) OPTSTAT is a convex optimization problem satisfying Slater's condition.*
(b) Let $\left(\left(\left(q_i^{\pi,1}(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^{\pi,1}(c)\right)_{c\in\mathcal{C}}\right)$ *and* $\left(\left(\left(q_i^{\pi,2}(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^{\pi,2}(c)\right)_{c\in\mathcal{C}}\right)$ *denote two optimal solutions to OPTSTAT. Then,* $q_i^{\pi,1}(f_i, l_i) = q_i^{\pi,2}(f_i, l_i)$ *for each* $(f_i, l_i) \in \mathcal{FL}_i$ *for each* $i \in \mathcal{N}$*.*
(c) Let $\left(\left(\left(q_i^{\pi,1}(f_i, l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^{\pi,1}(c)\right)_{c\in\mathcal{C}}\right)$ *denote an optimal solution to OPTSTAT. Then, given* $(f_i, l_i) \in \mathcal{FL}_i$, $q_i^{\pi,1}(f_i, l_i) = q_i^{\pi,1}(f_i, l_i')$ *for each* l_i' *such that* $(f_i, l_i') \in \mathcal{FL}_i$ *and for each* $i \in \mathcal{N}$*, i.e., quality choices for any two segments with the same QR tradeoff are the same irrespective of their lengths.*
(d) $\lim_{\delta\to 0}$ *OPTSTATVAL*$_\delta$*=OPTSTATVAL.*

Proof. Convexity properties of the objective and constraint functions of OPTSTAT are easy to establish using the convexity of the functions in $\mathcal{C}$ and $\cup_{i\in\mathcal{N}}\mathcal{F}_i$ once we establish convexity of $\mathrm{Var}\left(q_i\left(F_i^\pi, L_i^\pi\right)\right)$. This can be done using arguments similar to those in Lemma 3 (a) of [26], and we can show that OPTSTAT is a convex optimization problem. Using Assumption-SF, we can show that it also satisfies Slater's condition.

Proofs for part (b) is similar to that for Lemma 3 (b) in [26].

From (a), we can conclude that the KKT conditions are necessary and sufficient for optimality. Thus, there exist non-negative constants $\left(b_i^{\pi,1}\right)_{i\in\mathcal{N}}$, $\left(d_i^{\pi,1}\right)_{i\in\mathcal{N}}$, $\left(\left(\gamma_i^{\pi,1}(f_i,l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}$ and $\left(\left(\overline{\gamma}_i^{\pi,1}(f_i,l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}$, such that

$$\begin{aligned}
\left(U_i^E\right)'\left(m_i^{\pi,1}-U_i^V\left(v_i^{\pi,1}\right)\right)\left(1-2\left(U_i^V\right)'\left(v_i^{\pi,1}\right)\left(q_i^{\pi,1}(f_i,l_i)-m_i^{\pi,1}\right)\right)&\\
+\gamma_i^{\pi,1}(f_i,l_i)-\overline{\gamma}_i^{\pi,1}(f_i,l_i)-\frac{b_i^{\pi,1}}{\left(1+\overline{\beta}_i\right)}(f_i)'\left(q_i^{\pi,1}(f_i,l_i)\right)&\\
-p_i^d\frac{d_i^{\pi,1}}{\overline{p}_i}(f_i)'\left(q_i^{\pi,1}(f_i,l_i)\right) &= 0\ \forall\ (f_i,l_i)\in\mathcal{FL}_i,\ \forall\ i\in\mathcal{N}, && (3.47)\\
\gamma_i^{\pi,1}(f_i,l_i)\,q_i^{\pi,1}(f_i,l_i) &= 0,\ \forall\ (f_i,l_i)\in\mathcal{FL}_i,\ \forall\ i\in\mathcal{N}, && (3.48)\\
\overline{\gamma}_i^{\pi,1}(f_i,l_i)\left(q_{\max}-q_i^{\pi,1}(f_i,l_i)\right) &= 0,\ \forall\ (f_i,l_i)\in\mathcal{FL}_i,\ \forall\ i\in\mathcal{N}. && (3.49)
\end{aligned}$$

where for each $i\in\mathcal{N}$,

$$\begin{aligned}
m_i^{\pi,1} &= \frac{\mathbb{E}\left[L_i^\pi q_i^{\pi,1}\left(F_i^\pi,L_i^\pi\right)\right]}{\mathbb{E}\left[L_i^\pi\right]},\\
v_i^{\pi,1} &= \mathrm{Var}\left(q_i^{\pi,1}\left(F_i^\pi,L_i^\pi\right)\right).
\end{aligned}$$

Using (3.47), (3.48) and (3.49), we can conclude that, given $(f_i,l_i)\in\mathcal{FL}_i$, $q_i^{\pi,1}(f_i,l_i)$ is an optimal solution to

$$\begin{aligned}
\max_{0\le q\le q_{\max}} \left(U_i^E\right)'\left(m_i^{\pi,1}-U_i^V\left(v_i^{\pi,1}\right)\right)\left(q-\left(U_i^V\right)'\left(v_i^{\pi,1}\right)\left(q-m_i^{\pi,1}\right)^2\right)&\\
-\frac{b_i^{\pi,1}}{\left(1+\overline{\beta}_i\right)}f_i(q)-p_i^d\frac{d_i^{\pi,1}}{\overline{p}_i}f_i(q),&
\end{aligned}$$

Using the above observation, we have that $q_i^{\pi,1}(f_i,l_i)$ is the unique optimal solution to $\mathrm{QNOVA}_i\left(\left(m_i^{\pi,1},m_i^{\pi,1},v_i^{\pi,1},\left(h_i^B\right)^{-1}\left(b_i^{\pi,1}\right),\left(h_i^D\right)^{-1}\left(d_i^{\pi,1}\right)\right),f_i\right)$ (where the uniqueness is due to the strict concavity of the objective) which is independent of l_i and part (c) follows.

Part (d) is a result related to the continuity of optimal value of OPTSTAT$_\delta$ and this follows from Theorem 2.1 in [16] which provides sufficient conditions for verifying continuity of the optimal value. We can verify that the conditions given in Theorem 2.1 by noting that the objective function $\phi_\pi(.)$ (defined in (3.37)) of OPTSTAT is a continuous function (which follows from the continuity of the functions Mean$(.)$, Var$(.)$ and $U_i^V(.)$), and by establishing the upper semicontinuity, lower semicontinuity and compactness at $\delta = 0$ of the feasible region $\mathcal{QR}_\delta$ of OPTSTAT$_\delta$ defined below

$$\begin{aligned}\mathcal{QR}_\delta \quad = \quad & \Big\{ \Big(\Big((q_i(f_i, l_i))_{(f_i,l_i)\in\mathcal{FL}_i} \Big)_{i\in\mathcal{N}}, (\mathbf{r}(c))_{c\in\mathcal{C}} \Big) : \\ & \Big(\Big((q_i(f_i, l_i))_{(f_i,l_i)\in\mathcal{FL}_i} \Big)_{i\in\mathcal{N}}, (\mathbf{r}(c))_{c\in\mathcal{C}} \Big) \text{ satisfies the constraints} \\ & (3.39), (3.40), (3.41), (3.42), (3.45) \text{ and } (3.46) \Big\}\end{aligned}$$

The compactness of $\mathcal{QR}_\delta$ follows from the boundedness of the set, and the continuity of the functions associated with the constraints (3.39), (3.40), (3.41), (3.42), (3.45) and (3.46).

Proof of upper semicontinuity: We say that $\mathcal{QR}_\delta$ is upper semicontinuous at $\delta = 0$ if for each open set $\mathcal{QR}$ such that $\mathcal{QR}_0 \subset \mathcal{QR}$, there is some $\overline{\delta} > 0$ such that $\mathcal{QR}_\delta \subset \mathcal{QR}$ for each $\delta \in [-\overline{\delta}, \overline{\delta}]$. Since $\mathcal{QR}_0$ is a compact set and $\mathcal{QR}$ is an open set, we can find $\delta_0 > 0$ such that δ_0 expansion of $\mathcal{QR}_0$ is a subset of $\mathcal{QR}$ (since each point in $\mathcal{QR}_0$ is an interior point of $\mathcal{QR}$, we can obtain an open cover of the compact set $\mathcal{QR}_0$ comprising the union of neighborhoods of positive radii centered points in $\mathcal{QR}_0$, and then we can obtain $\delta_0 > 0$ as the minimum radius of neighborhoods associated with a finite subcover).

Next, we show that the distance between any point $\Big(\Big((q_i(f_i, l_i))_{(f_i,l_i)\in\mathcal{FL}_i} \Big)_{i\in\mathcal{N}}, (\mathbf{r}(c))_{c\in\mathcal{C}} \Big) \in \mathcal{QR}_\delta$ and the set $\mathcal{QR}_0$ can be made as close to zero as desired by picking δ small enough. Since this is trivial for $\delta \leq 0$, we focus on $\delta > 0$. Consider $\left(\left(q_i'\left(f_i, l_i\right)\right)_{(f_i, l_i) \in \mathcal{FL}_i}\right)_{i \in \mathcal{N}}$ defined as follows

$$q_i'\left(f_i, l_i\right) = \frac{q_i\left(f_i, l_i\right)}{1 + \xi\delta} \ \forall \ \left(f_i, l_i\right) \in \mathcal{FL}_i, \ i \in \mathcal{N}, \tag{3.50}$$

where

$$\xi = \max_{i \in \mathcal{N}} \max\left(\frac{1}{\left(1 - p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi(0)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]}\right)}, \frac{1}{\left(1 + \frac{\mathbb{E}\left[r_i(C^\pi)\right]}{\tau_{slot}} - \frac{\mathbb{E}\left[L_i^\pi F_i^\pi(0)\right]}{(1+\overline{\beta}_i)\mathbb{E}\left[L_i^\pi\right]}\right)}\right).$$

We can show that $\xi \geq 0$ using Assumption-SF, and hence

$$q_i'\left(f_i, l_i\right) \leq q_i\left(f_i, l_i\right) \ \forall \ \left(f_i, l_i\right) \in \mathcal{FL}_i, \ i \in \mathcal{N}. \tag{3.51}$$

Now, consider the following expression in the left hand side of the cost constraint (3.45) in OPTSTAT$_\delta$ evaluated at $\left(\left(\left(q_i'\left(f_i, l_i\right)\right)_{(f_i, l_i) \in \mathcal{FL}_i}\right)_{i \in \mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c \in \mathcal{C}}\right)$

$$\begin{aligned}
p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i'\left(F_i^\pi, L_i^\pi\right)\right)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]} &\leq \frac{1}{1+\xi\delta} p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi, L_i^\pi\right)\right)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]} \\
&\quad + \frac{\xi\delta}{1+\xi\delta} p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]} \\
&\leq \frac{1+\delta}{1+\xi\delta} + \frac{\xi\delta}{1+\xi\delta} p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]} \\
&= \frac{1 + \delta + \xi\delta p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi(0)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]}}{1+\xi\delta} \\
&\leq 1,
\end{aligned} \tag{3.52}$$

where the first inequality above follows from the convexity of functions in $\mathcal{F}_i$, and the second inequality is due to the fact that $\left(\left(\left(q_i\left(f_i, l_i\right)\right)_{(f_i, l_i) \in \mathcal{FL}_i}\right)_{i \in \mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c \in \mathcal{C}}\right) \in \mathcal{QR}_\delta$ (and hence satisfies (3.45)). Note that the expression in the third line is less than or equal to one if $\delta + \xi\delta p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi(0)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]} \leq \xi\delta$, and this holds due to our choice of ξ which ensures that $\xi \geq 1 + \xi p_i^d \frac{\mathbb{E}\left[L_i^\pi F_i^\pi(0)\right]}{\overline{p}_i \mathbb{E}\left[L_i^\pi\right]}$ for each $i \in \mathcal{N}$. Next, consider the following expression from the rebuffering constraint (3.46) in OPTSTAT$_\delta$ evaluated at $\left(\left(\left(q_i'\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)$

$$\begin{aligned}
\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i'\left(F_i^\pi, L_i^\pi\right)\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} &\leq \frac{1}{1+\xi\delta}\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi, L_i^\pi\right)\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} + \frac{\xi\delta}{1+\xi\delta}\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} \\
&\leq \frac{1+\delta+\frac{\mathbb{E}\left[r_i(C^\pi)\right]}{\tau_{slot}}}{1+\xi\delta} + \frac{\xi\delta}{1+\xi\delta}\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} \\
&= \frac{1+\delta+\frac{\mathbb{E}\left[r_i(C^\pi)\right]}{\tau_{slot}} + \xi\delta\frac{\mathbb{E}\left[L_i^\pi F_i^\pi(0)\right]}{(1+\overline{\beta}_i)\mathbb{E}\left[L_i^\pi\right]}}{1+\xi\delta} \\
&\leq 1, \qquad (3.53)
\end{aligned}$$

where the above inequalities follow from arguments similar to those made to obtain (3.52). Using the fact that $\delta > 0$ and $\left(\left(\left(q_i\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right) \in \mathcal{QR}_\delta$, and using the inequalities (3.51), (3.52) and (3.53), we can conclude that $\left(\left(\left(q_i'\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right) \in \mathcal{QR}_0$. Then, by using the expression for $\left(\left(q_i'\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}$ in (3.50), we can conclude that the distance between $\left(\left(\left(q_i\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)$ and $\left(\left(\left(q_i'\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)$ is $O(\delta)$. Hence, the distance between any $\left(\left(\left(q_i\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right) \in \mathcal{QR}_\delta$ and the set $\mathcal{QR}_0$ goes to zero as δ goes to zero.

Hence, we can always $\overline{\delta} > 0$ such that for each $\delta \in [-\overline{\delta}, \overline{\delta}]$, $\mathcal{QR}_\delta$ is contained in a δ_0 expansion of $\mathcal{QR}_0$ which in turn is a subset of the open set $\mathcal{QR}$ containing $\mathcal{QR}_0$. Hence, for each $\delta \in [-\overline{\delta}, \overline{\delta}]$, $\mathcal{QR}_\delta \subset \mathcal{QR}$ for each $\delta \in [-\overline{\delta}, \overline{\delta}]$. Thus, $\mathcal{QR}_\delta$ is upper semicontinuous at $\delta = 0$.

Proof of lower semicontinuity: We show that $\mathcal{QR}_\delta$ is lower semicontinuous at $\delta = 0$ by showing that it is open at $\delta = 0$. $\mathcal{QR}_\delta$ is open at $\delta = 0$ if for any sequence $(\delta_n)_{n\geq 1}$ converging to 0 and $\left(\left(\left(q_i\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right) \in \mathcal{QR}_0$, we can find a sequence $\left(\left(\left(q_i^{(n)}\left(f_i, l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^{(n)}\left(c\right)\right)_{c\in\mathcal{C}}\right) \in \mathcal{QR}_{\delta_n}$

that converges to $\left(\left(\left(q_i\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},\left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)$. We can obtain the desired sequence by setting $\mathbf{r}^{(n)}\left(c\right)=\mathbf{r}\left(c\right)$ for each $c\in\mathcal{C}$, and

$$q_i^{(n)}\left(f_i,l_i\right)=\min(1+\xi\delta_n,1)q_i\left(f_i,l_i\right)\ \forall\ \left(f_i,l_i\right)\in\mathcal{FL}_i,\ i\in\mathcal{N}. \tag{3.54}$$

Next, we verify that $\left(\left(\left(q_i^{(n)}\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},\left(\mathbf{r}^{(n)}\left(c\right)\right)_{c\in\mathcal{C}}\right)\in\mathcal{QR}_{\delta_n}$ for each n. This is clear for $\delta_n\geq 0$, and hence we restrict our attention to $\delta_n<0$. Consider the following expression in the left hand side of the cost constraint (3.45) in OPTSTAT$_\delta$ evaluated at $\left(\left(\left(q_i'\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},\left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)$

$$\begin{aligned}
p_i^d\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i^{(n)}\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\overline{p}_i\mathbb{E}\left[L_i^\pi\right]} &\leq \left(1+\xi\delta\right)p_i^d\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\overline{p}_i\mathbb{E}\left[L_i^\pi\right]}\\
&\qquad -\xi\delta p_i^d\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\overline{p}_i\mathbb{E}\left[L_i^\pi\right]}\\
&\leq 1+\xi\delta-\xi\delta p_i^d\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\overline{p}_i\mathbb{E}\left[L_i^\pi\right]}\\
&\leq 1,
\end{aligned} \tag{3.55}$$

where the above inequalities can be shown using arguments similar to those made to obtain (3.52) and (3.53). Similarly, we can show that

$$\begin{aligned}
\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i^{(n)}\left(F_i^\pi,L_i^\pi\right)\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]} &\leq \left(1+\xi\delta\right)\left(1+\frac{\mathbb{E}\left[r_i\left(C^\pi\right)\right]}{\tau_{slot}}\right)-\xi\delta\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(0\right)\right]}{\left(1+\overline{\beta}_i\right)\mathbb{E}\left[L_i^\pi\right]}\\
&\leq 1.
\end{aligned} \tag{3.56}$$

Using the fact that $\left(\left(\left(q_i\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},\left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)\in\mathcal{QR}_0$, (3.54), (3.55) and (3.56), we can conclude that $\left(\left(\left(q_i^{(n)}\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},\left(\mathbf{r}^{(n)}\left(c\right)\right)_{c\in\mathcal{C}}\right)\in\mathcal{QR}_{\delta_n}$ for each n, and the associated sequence converges to $\left(\left(\left(q_i\left(f_i,l_i\right)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}},\left(\mathbf{r}\left(c\right)\right)_{c\in\mathcal{C}}\right)$. Thus, we have showed that $\mathcal{QR}_\delta$ is open at $\delta=0$, and is thus lower semicontinuous at $\delta=0$

□

We let $\left(\left(q_i^\pi(f)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$ denote the optimal quality choices associated with different quality rate tradeoffs. Note that we have dropped the dependence of the optimal quality choices on segment length based on the observation in Lemma 3.4 (c). Let $\left(\left(\left(q_i^\pi(f)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^\pi(c)\right)_{c\in\mathcal{C}}\right)$ be an optimal solution to OPTSTAT, and let $\mathbf{b}^\pi$ and $\mathbf{d}^\pi$ denote the associated Lagrange multipliers for the constraints (3.43) and (3.44) respectively. Using the above result, we can conclude that the KKT conditions are necessary and sufficient for optimality. Hence, there exist non-negative constants (referred to in the sequel as Lagrange multipliers associated with the optimal solution) $(\chi^\pi(c))_{c\in\mathcal{C}}$, $\left((\gamma_i^\pi(f))_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$, $\left((\overline{\gamma}_i(f)^\pi)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$, $(\boldsymbol{\omega}^\pi(c))_{c\in\mathcal{C}}$, $\mathbf{d}^\pi$ and $\mathbf{b}^\pi$ such that

$$\begin{aligned}
\left(U_i^E\right)' \left(m_i^\pi - U_i^V\left(v_i^\pi\right)\right)\left(1 - 2\left(U_i^V\right)'\left(v_i^\pi\right)\left(q_i^\pi(f) - m_i^\pi\right)\right) + \gamma_i^\pi(f) - \overline{\gamma}_i^\pi(f) & \\
- \frac{b_i^\pi}{\left(1+\overline{\beta}_i\right)}(f)'\left(q_i^\pi(f)\right) - p_i^d \frac{d_i^\pi}{\overline{p}_i}(f)'\left(q_i^\pi(f)\right) &= 0,\ \forall\ f\in\mathcal{F}_i, \forall\ i\in\mathcal{N}, \quad (3.57)\\
-\chi^\pi(c)\frac{\partial c\left(\mathbf{r}^\pi(c)\right)}{\partial r_i} + \frac{b_i^\pi}{\tau_{slot}} + \omega_i^\pi(c) &= 0,\ \forall\ c\in\mathcal{C}, \forall\ i\in\mathcal{N}, \quad (3.58)
\end{aligned}$$

$$\begin{aligned}
\gamma_i^\pi(f) q_i^\pi(f) &= 0,\ \forall\ f\in\mathcal{F}_i, \forall\ i\in\mathcal{N}, && (3.59)\\
\overline{\gamma}_i^\pi(f)\left(q_{\max} - q_i^\pi(f)\right) &= 0,\ \forall\ f\in\mathcal{F}_i, \forall\ i\in\mathcal{N}, && (3.60)\\
\chi^\pi(c) c\left(\mathbf{r}^\pi(c)\right) &= 0, && (3.61)\\
\frac{\omega_i^\pi(c)}{K_S}\left(r_i^\pi(c) - r_{i,\min}\right) &= 0,\ \forall\ c\in\mathcal{C}, \forall\ i\in\mathcal{N}, && (3.62)\\
d_i^\pi\left(1 - \frac{p_i^d\sigma_i^\pi}{\overline{p}_i}\right) &= 0\ \forall\ i\in\mathcal{N}, && (3.63)\\
b_i^\pi\left(\frac{\sigma_i^\pi}{\left(1+\overline{\beta}_i\right)} - \frac{\rho_i^\pi}{\tau_{slot}}\right) &= 0\ \forall\ i\in\mathcal{N}. && (3.64)
\end{aligned}$$

where for each $i \in \mathcal{N}$,

$$
\begin{aligned}
m_i^\pi &= \frac{\mathbb{E}\left[L_i^\pi q_i^\pi\left(F_i^\pi\right)\right]}{\mathbb{E}\left[L_i^\pi\right]}, && (3.65)\\
v_i^\pi &= \operatorname{Var}\left(q_i^\pi\left(F_i^\pi\right)\right), && (3.66)\\
\sigma_i^\pi &= \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i^\pi\left(F_i^\pi\right)\right)\right]}{\mathbb{E}\left[L_i^\pi\right]}, && (3.67)\\
\lambda_i^\pi &= \mathbb{E}\left[L_i^\pi\right]. && (3.68)
\end{aligned}
$$

Thus m_i^π, v_i^π and σ_i^π are the (statistical) mean quality, variance in quality and mean segment size for video client i associated with optimal solution to OPTSTAT. Also, let

$$
\begin{aligned}
\mathcal{X}^\pi = \Big\{ & \left(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi, \mathbf{d}^\pi\right) : \text{there is an optimal solution} && (3.69)\\
& \left(\left(\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^\pi\left(c\right)\right)_{c\in\mathcal{C}}\right) \text{ to OPTSTAT with}\\
& \rho_i^\pi = \mathbb{E}\left[r_i^\pi\left(C^\pi\right)\right] \text{ for each } i \in \mathcal{N}, \text{ and with}\\
& \mathbf{b}^\pi \text{ and } \mathbf{d}^\pi \text{ as the associated optimal Lagrange multipliers}\\
& \text{for the constraints (3.43) and (3.44) respectively}\Big\}.
\end{aligned}
$$

In the next result, we present three useful properties of the optimal solution to OPTSTAT. The result in part (a) below provides a video client level optimality result which essentially suggests that we can decouple the quality adaptation of the video clients. It states that the component $\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}$ of the optimal solution to OPTSTAT associated with video client $i \in \mathcal{N}$ is itself an optimal solution to an optimization problem which can be solved by the video client i. This result hints at the possibility of distributing the task of quality adaptation across the video clients so that each video client manages its own adaptation. The result in part (b) points out that we only need to know a few parameters (specifically, the optimal Lagrange multipliers associated with the rebuffering constraints) associated with the quality adaptation to carry out optimal resource allocation. This suggests that we could potentially decouple the task of optimal resource allocation from quality adaptation. Part (c) states that that when NOVA parameter $\boldsymbol{\theta}_{i,s}$ of video client i is in the set $\mathcal{H}_i^*$ defined below

$$\mathcal{H}_i^{*} := \left\{ \left(m_i^{\pi}, m_i^{\pi}, v_i^{\pi}, \left(h_i^{B}\right)^{-1}\left(b_i^{\pi}\right), \left(h_i^{D}\right)^{-1}\left(d_i^{\pi}\right) \right) : \left(\boldsymbol{\rho}^{\pi}, \mathbf{b}^{\pi}, \mathbf{d}^{\pi}\right) \in \mathcal{X}^{\pi} \right\}, \tag{3.70}$$

we can obtain optimal quality choices for OPTSTAT by using NOVA.

Lemma 3.5. *For parts (a) and (b) of this result, suppose* $\left(\boldsymbol{\rho}^{\pi}, \mathbf{b}^{\pi}, \mathbf{d}^{\pi}\right) \in \mathcal{X}^{\pi}$*, and let* $\left(\left(\left(q_i^{\pi}(f)\right)_{f \in \mathcal{F}_i} \right)_{i \in \mathcal{N}}, \left(\mathbf{r}^{\pi}(c)\right)_{c \in \mathcal{C}} \right)$ *be the associated optimal solution.*
(a) For each $i \in \mathcal{N}$*,* $\left(q_i^{\pi}(f)\right)_{f \in \mathcal{F}_i}$ *is the unique optimal solution to the following optimization problem*

$$\begin{aligned}
\max_{\left((q_i(f))_{f \in \mathcal{F}_i} \right)} \; & U_i^{E}\left(\frac{\mathbb{E}\left[L_i^{\pi} q_i\left(F_i^{\pi}\right)\right]}{\mathbb{E}\left[L_i^{\pi}\right]} - U_i^{V}\left(Var\left(q_i\left(F_i^{\pi}\right)\right)\right) \right) \\
& - \sum_{i \in \mathcal{N}} d_i^{\pi} \left(\frac{p_i^{d}}{\overline{p}_i} \right) \left(\frac{\mathbb{E}\left[L_i^{\pi} F_i^{\pi}\left(q_i\left(F_i^{\pi}\right)\right)\right]}{\mathbb{E}\left[L_i^{\pi}\right]} \right) - \sum_{i \in \mathcal{N}} \frac{b_i^{\pi}}{\left(1 + \overline{\beta}_i\right)} \left(\frac{\mathbb{E}\left[L_i^{\pi} F_i^{\pi}\left(q_i\left(F_i^{\pi}\right)\right)\right]}{\mathbb{E}\left[L_i^{\pi}\right]} \right), \\
q_i(f) \;\; \geq \;\; & 0, \; \forall \; f \in \mathcal{F}_i, \\
q_i(f) \;\; \leq \;\; & q_{\max}, \; \forall \; f \in \mathcal{F}_i.
\end{aligned}$$

(b) $\left(\mathbf{r}^{\pi}(c)\right)_{c \in \mathcal{C}}$ *is an optimal solution to the following optimization problem*

$$\begin{aligned}
& \mathbb{E}\left[\sum_{i \in \mathcal{N}} b_i^{\pi} r_i\left(C^{\pi}\right) \right], \\
s.t. \;\; c\left(\mathbf{r}(c)\right)) \;\; & \leq \;\; 0, \; \forall \; c \in \mathcal{C}, \\
r_i(c) \;\; & \geq \;\; r_{i,\min}, \; \forall \; c \in \mathcal{C}, \forall \; i \in \mathcal{N}.
\end{aligned}$$

(c) The following holds for each $i \in \mathcal{N}$: If $\boldsymbol{\theta}_i^\pi \in \mathcal{H}_i^*$, then $q_i^*\left(\boldsymbol{\theta}_i^\pi, f\right) = q_i^\pi\left(f\right)$ for each $f \in \mathcal{F}_i$.

Proof. As with the case of OPTSTAT, we can show that KKT conditions are necessary and sufficient for optimality for the optimization problem considered in part (a). Now, the result follows by using (3.57), (3.59) and (3.60) to conclude that $\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}$ satisfies these conditions. Proof of part (b) is similar to that of (a), and can be completed by using the fact that $\left(\mathbf{r}^\pi\left(c\right)\right)_{c\in\mathcal{C}}$ satisfies (3.58), (3.61) and (3.62).

Using the necessary optimality conditions for OPTSTAT given in (3.57), (3.59) and (3.60), we can show that $q_i^\pi\left(f\right)$ satisfies the sufficient optimality conditions (3.32)-(3.34) for $\text{QNOVA}_i\left(\boldsymbol{\theta}_i^\pi, f\right)$ (following an approach similar to that used in the proof of part (c) of Lemma 3.4). Then part (c) follows from the fact that $\text{QNOVA}_i\left(\boldsymbol{\theta}_i^\pi, f\right)$ has a unique optimal solution.

□

We use the observation in part (c) and properties of OPTSTAT to prove the next result which is the main result for this subsection and is an important intermediate result used in the proof of main optimality of NOVA given in Theorem 3.1. The result states that the performance of NOVA with its parameters $\boldsymbol{\theta}_{i,s}$ picked from the set $\mathcal{H}_i^*$ for each $i \in \mathcal{N}$ is asymptotically optimal. Further, this result suggests that we can prove Theorem 3.1 if we can show that the updates (3.21)-(3.26) of NOVA guide the parameters $\left(\boldsymbol{\theta}_{i,s}\right)_{s\geq 1}$ of video client i to $\mathcal{H}_i^*$ for each video client $i \in \mathcal{N}$. This motivates the study of convergence behavior of NOVA which is the main focus of the rest of this section.

Theorem 3.2. *Suppose $\boldsymbol{\theta}_i^\pi \in \mathcal{H}_i^*$ for each $i \in \mathcal{N}$. Then, for almost all sample paths*

$$\lim_{S\to\infty}\left(\phi_S\left(\left(\left(q_i^*\left(\boldsymbol{\theta}_i^\pi, f_{i,s}\right)\right)_{i\in\mathcal{N}}\right)_{1\leq s\leq S}\right) - \phi_S^{opt}\right) = 0.$$

Proof. For a fixed S, consider an optimal solution $\left(\left(\mathbf{q}^S\right)_{1:S}, \left(\mathbf{r}^S\right)_{1:K_S}\right)$ to OPT(S). Without loss of generality (we prove this below), we assume that the optimal solution $\left(\left(\mathbf{q}^S\right)_{1:S}, \left(\mathbf{r}^S\right)_{1:K_S}\right)$ satisfies the following two conditions:
(a) $q^S_{i,s_1} = q^S_{i,s_2}$ for any two segments s_1 and s_2 such that $f_{i,s_1} = f_{i,s_2}$ and $l_{i,s_1} = l_{i,s_2}$.
(b) $\mathbf{r}^S_{k_1} = \mathbf{r}^S_{k_2}$ for any two slots k_1 and k_2 such that $c_{k_1} = c_{k_2}$.
We show that we can always find an optimal solution satisfying these conditions. Consider an optimal solution $\left(\left(\mathbf{q}^{0,S}\right)_{1:S}, \left(\mathbf{r}^{0,S}\right)_{1:K_S}\right)$ to OPT(S) that does not satisfy the conditions. We can obtain another optimal solution $\left(\left(\mathbf{q}^{1,S}\right)_{1:S}, \left(\mathbf{r}^{1,S}\right)_{1:K_S}\right)$ to OPT(S) satisfying this condition by letting

$$\begin{aligned} q^{1,S}_{i,s'} &= \frac{\sum_{s=1}^{S} I\left(f_{i,s} = f_{i,s'}, l_{i,s} = l_{i,s'}\right) q^{0,S}_{i,s}}{\sum_{s=1}^{S} I\left(f_{i,s} = f_{i,s'}, l_{i,s} = l_{i,s'}\right)}, \ \forall\, 1 \le s' \le S, \\ r^{1,S}_{i,k'} &= \frac{\sum_{k=1}^{K_S} I\left(c_k = c_{k'}\right) r^{0,S}_{i,k}}{\sum_{k=1}^{K_S} I\left(c_k = c_{k'}\right)}, \ \forall\, 1 \le k' \le K_S. \end{aligned}$$

It is clear that $\left(\left(\mathbf{q}^{1,S}\right)_{1:S}, \left(\mathbf{r}^{1,S}\right)_{1:K_S}\right)$ satisfies the conditions (a) and (b). Further, using the structure of the optimization problem OPT(S), we can show that $\left(\left(\mathbf{q}^{1,S}\right)_{1:S}, \left(\mathbf{r}^{1,S}\right)_{1:K_S}\right)$ is also an optimal solution to OPT(S).

Now, we return to the proof of the main result and consider an optimal solution $\left(\left(\mathbf{q}^S\right)_{1:S}, \left(\mathbf{r}^S\right)_{1:K_S}\right)$ to OPT(S) satisfying the conditions (a) and (b) so that the component $q^S_{i,s}$ in the optimal solution assocaited with quality adaptation for segment s of video client i depends only on (f, l). Hence, we can obtain a function $q^S_i(f, l)$ for $(f, l) \in \mathcal{FL}_i$, such that $q^S_i(f, l)$ denotes the quality associated with this optimal solution for a segment s with QR tradeoff $f_{i,s} = f$ and length $l_{i,s} = l$. Similarly, we can obtain a function $r^S_i(c)$ for $c \in \mathcal{C}$ such that $r^S_i(c)$ denotes the resource allocation associated with the optimal solution for a slot k with allocation constraint $c_k = c$. Let $\pi_i^{\mathcal{F},\mathcal{L},S}(f_i, l_i) = \frac{\sum_{s=1}^{S} I(f_{i,s}=f_i,\ l_{i,s}=l_i)}{S}$ for $(f, l) \in \mathcal{FL}_i$ be the empirical distribution for the occurrence of segments with QR tradeoff f and length l, and $\pi^{\mathcal{C},S}(c) = \frac{\sum_{k=1}^{K_S} I(c_k=c)}{K_S}$ for $c \in \mathcal{C}$ be the empirical distribution for the occurrence of allocation constraint c.

The mean quality for video client i corresponding to the optimal solution $\left(\left(\mathbf{q}^S\right)_{1:S}, \left(\mathbf{r}^S\right)_{1:K_S}\right)$ is given by

$$\begin{aligned} m_i^S\left(q_i^S\right) &= \frac{\sum_{s=1}^{S} l_{i,s} q_{i,s}^S}{\sum_{s=1}^{S} l_{i,s}} = \frac{\sum_{(f_i,l_i)\in\mathcal{FL}_i} \pi_i^{\mathcal{F},\mathcal{L},S}(f_i,l_i)\, l_i q_i^S(f_i,l_i)}{\sum_{(f_i,l_i)\in\mathcal{FL}_i} \pi_i^{\mathcal{F},\mathcal{L},S}(f_i,l_i)\, l_i} \\ &= \frac{\sum_{(f_i,l_i)\in\mathcal{FL}_i} \pi_i^{\mathcal{F},\mathcal{L}}(f_i,l_i)\, l_i q_i^S(f_i,l_i)}{\sum_{(f_i,l_i)\in\mathcal{FL}_i} \pi_i^{\mathcal{F},\mathcal{L}}(f_i,l_i)\, l_i} + \delta_m(S) \end{aligned}$$

where $\delta_m(S)$ is a function such that $\lim_{S\to\infty} \delta_m(S) = 0$ a.s. (in this proof, 'a.s.' stands for 'for almost all sample paths'). This limiting behavior of $\delta_m(S)$ follows from the boundedness of the terms involved, and the fact that $\left(L_{i,s}, F_{i,s}\right)_{s\geq 0}$ is a stationary ergodic process as a result of which $\lim_{S\to\infty}\left|\pi_i^{\mathcal{F},\mathcal{L},S}(f_i,l_i) - \pi_i^{\mathcal{F},\mathcal{L}}(f_i,l_i)\right| = 0$ a.s. for each $(f,l) \in \mathcal{FL}_i$, i.e., the empirical distribution converges to the stationary distribution. Recall that $\left(\pi_i^{\mathcal{F},\mathcal{L}}(f_i,l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}$ is the marginal distribution associated with the stationary ergodic process $\left(F_{i,s}, L_{i,s}\right)_{s\geq 0}$. Using similar calculations, we can obtain a function $\delta_{e,1}(S)$ satisfying $\lim_{S\to\infty} \delta_{e,1}(S) = 0$ a.s. such that the optimal value for OPT(S), i.e. $\phi_S^{opt} = \phi_S\left(\left(\mathbf{q}^S\right)_{1:S}\right)$ can be expressed as

$$\phi_S^{opt} = \sum_{i\in\mathcal{N}} U_i^E\left(\text{Mean}\left(q_i^S\left(F_i^\pi, L_i^\pi\right)\right) - U_i^V\left(\text{Var}\left(q_i^S\left(F_i^\pi, L_i^\pi\right)\right)\right)\right) + \delta_{e,1}(S).$$

where the first term on the right hand side is equal to the objective function of OPTSTAT (given in (3.37)) evaluated at $\left(\left(q_i^S(f_i,l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}$.

Again using the fact that $\lim_{S\to\infty}\left|\pi_i^{\mathcal{F},\mathcal{L},S}(f_i,l_i)_i - \pi^{\mathcal{F},\mathcal{L}}(f_i,l_i)\right| = 0$ a.s. and $\lim_{S\to\infty}\left|\pi^{\mathcal{C},S}(c) - \pi^{\mathcal{C}}(c)\right| = 0$ a.s., together with arguments similar to those above, we can show that $\left(\left(\left(q_i^S(f_i,l_i)\right)_{(f_i,l_i)\in\mathcal{FL}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}'_S(c)\right)_{c\in\mathcal{C}}\right)$ is a feasible solution to the optimization problem OPTSTAT$_{\delta_{e,2}(S)}$ (i.e., optimization problem OPTSTAT with constraints loosened by $\delta_{e,2}(S)$) for an appropriately chosen function $\delta_{e,2}(S)$ satisfying $\lim_{S\to\infty}\delta_{e,2}(S)=0$ a.s.. Hence,

$$\phi_S^{opt} \leq \text{OPTSTATVAL}_{\delta_{e,2}(S)} + \delta_{e,1}(S) \tag{3.71}$$

From Lemma 3.4 (d), we have that $\lim_{\delta\to 0}$ OPTSTATVAL$_\delta$=OPTSTATVAL. Using Lemma 3.5 (c) and the fact that $(F_{i,s}, L_{i,s})_{s\geq 0}$ is a stationary ergodic process for each $i \in \mathcal{N}$, we have

$$\lim_{S\to\infty}\left(\phi_S\left(\left((q_i^*\left(\boldsymbol{\theta}_i^\pi, f_{i,s}\right))_{i\in\mathcal{N}}\right)_{1\leq s\leq S}\right) - \text{OPTSTATVAL}\right) = 0 \text{ a.s..}$$

Now, the result follows by using these two observations, (3.71) and the fact that

$$\phi_S\left(\left((q_i^*\left(\boldsymbol{\theta}_i^\pi, f_{i,s}\right))_{i\in\mathcal{N}}\right)_{1\leq s\leq S}\right) \leq \phi_S^{opt}.$$

□

3.5.2 An auxiliary differential inclusion related to NOVA

In the previous subsection, we stated Theorem 3.2 which suggests that we can prove the main optimality result for NOVA if we establish an appropriate convergence result for NOVA. In this subsection, we study an auxiliary differential inclusion which evolves according to average dynamics of NOVA. The main goal of this subsection is to study the convergence of the differential inclusion which in turn will help us obtain the desired convergence of parameters of NOVA in the next subsection.

For the rest of this section, we additionally consider the evolution of auxiliary parameters $(\sigma_{i,s_i})_{s_i\geq 1}$ and $(\rho_{i,k})_{k\geq 1}$ associated with NOVA which evolve according to update rules discussed next. We update the auxiliary parameter σ_{i,s_i} based on the quality q^*_{i,s_i+1} (shorthand for $q_i^*(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$, $\boldsymbol{\theta}_{i,s_i} = (m_{i,s_i}, \mu_{i,s_i}, v_{i,s_i}, b_{Q,i,s_i}, d_{i,s_i})$)

chosen by NOVA for (s_i+1)th segment of video client $i \in \mathcal{N}$ as follows:

$$\sigma_{i,s_i+1} = \sigma_{i,s_i} + \epsilon \left(\frac{l_{i,s_i+1} f_{i,s_i}\left(q^*_{i,s_i+1}\right)}{\lambda_{i,s_i}} - \sigma_{i,s_i} \right), \tag{3.72}$$

Thus, the auxiliary parameter σ_{i,s_i} tracks the mean segment size of the segments downloaded by video client $i \in \mathcal{N}$. We update the parameter $\boldsymbol{\rho}_k$ based on the resource allocation $\mathbf{r}^*_k \in \mathcal{R}^*(\mathbf{b}_k, c_k)$ in slot k as described below

$$\rho_{i,k+1} = \rho_{i,k} + \epsilon \left(r^*_{i,k} - \rho_{i,k}\right) \ \forall\ i \in \mathcal{N}. \tag{3.73}$$

Thus, the auxiliary parameter $\boldsymbol{\rho}_k$ tracks the mean resource allocation to video clients. Note that the auxiliary parameters $\sigma_{i,\cdot}$ and $\boldsymbol{\rho}_{\cdot}$ do not affect the allocation or adaptation in NOVA.

Next, let

$$\begin{aligned} \mathcal{H} = & \big\{ (\mathbf{m}, \boldsymbol{\mu}, \mathbf{v}, \mathbf{b}, \mathbf{d}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathbb{R}^{8N} : \text{ for each } i \in \mathcal{N}, \\ & 0 \le m_i, \mu_i \le q_{\max}, 0 \le v_i \le q^2_{\max},\ \underline{b} \le b_i \le \overline{b},\ \underline{d} \le d_i \le \overline{d}, \\ & l_{\min} \le \lambda_i \le l_{\max},\ l_{\min} f_{\min} \le \sigma_i \le l_{\max} f_{\max},\ r_{i,\min} \le \rho_i \le r_{\max} \big\}. \end{aligned} \tag{3.74}$$

Note that the parameters $(\mathbf{m}_s, \boldsymbol{\mu}_s, \mathbf{v}_s, \mathbf{b}_k, \mathbf{d}_s, \boldsymbol{\lambda}_s, \boldsymbol{\sigma}_s, \boldsymbol{\rho}_k)_{s,k}$ associated with NOVA remain in $\mathcal{H}$ (see Lemma 3.1). For each video client $i \in \mathcal{N}$, we use the variables $\widehat{m}_i(t)$, $\widehat{\mu}_i(t)$, $\widehat{v}_i(t)$, $\widehat{b}_i(t)$, $\widehat{d}_i(t)$, $\widehat{\lambda}_i(t)$, $\widehat{\sigma}_i(t)$ and $\widehat{\rho}_i(t)$ to track the average dynamics of the parameters m_{i,s_i}, μ_{i,s_i}, v_{i,s_i}, $b_{i,k}$, d_{i,s_i}, λ_{i,s_i}, σ_{i,s_i} and $\rho_{i,k}$ respectively associated with NOVA (and this is explained in detail in the sequel before Lemma 3.6). Let $\widehat{\boldsymbol{\Theta}}(t) = \left(\widehat{\mathbf{m}}(t), \widehat{\boldsymbol{\mu}}(t), \widehat{\mathbf{v}}(t), \widehat{\mathbf{b}}(t), \widehat{\mathbf{d}}(t), \widehat{\boldsymbol{\lambda}}(t), \widehat{\boldsymbol{\sigma}}(t), \widehat{\boldsymbol{\rho}}(t)\right) \in \mathcal{H}$ and $\widehat{\boldsymbol{\theta}}_i(t) = (\widehat{m}_i(t), \widehat{\mu}_i(t), \widehat{v}_i(t), \widehat{b}_i(t), \widehat{d}_i(t))$ for each $i \in \mathcal{N}$, i.e., $\widehat{\boldsymbol{\theta}}_i(t)$ includes the components in $\widehat{\boldsymbol{\Theta}}(t)$ that affect the quality adaptation of video client $i \in \mathcal{N}$.

The main focus of this subsection is the following differential inclusion which describes the evolution of $\left(\widehat{\boldsymbol{\Theta}}(t)\right)_{t\geq 0}$:

Auxiliary differential inclusion related to NOVA

$\widehat{\boldsymbol{\Theta}}(0) \in \mathcal{H}$ and for almost all $t \geq 0$ and each $i \in \mathcal{N}$,

$$\dot{\widehat{m}}_i(t) = \frac{\left(U_i^E\right)'\left(\widehat{\mu}_i(t) - U_i^V\left(\widehat{v}_i(t)\right)\right)\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} \left(\frac{E\left[L_i^\pi q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right]}{\widehat{\lambda}_i(t)} - \widehat{m}_i(t)\right), \tag{3.75}$$

$$\dot{\widehat{\mu}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}\left(\frac{E\left[L_i^\pi q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right]}{\widehat{\lambda}_i(t)} - \widehat{\mu}_i(t)\right), \tag{3.76}$$

$$\dot{\widehat{v}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}\left(\frac{E\left[L_i^\pi\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right) - \widehat{m}_i(t)\right)^2\right]}{\widehat{\lambda}_i(t)} - \widehat{v}_i(t)\right), \tag{3.77}$$

$$\dot{\widehat{b}}_i(t) = \frac{1}{\left(1+\overline{\beta}_i\right)} - \frac{E\left[L_i^\pi\right]}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} + \widehat{z}_i^b\left(\widehat{\boldsymbol{\Theta}}(t)\right), \tag{3.78}$$

$$\dot{\widehat{d}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}\left(\frac{p_i^d E\left[L_i^\pi F_i^\pi\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right)\right]}{\overline{p}_i} - \widehat{\lambda}_i(t)\right) + \widehat{z}_i^d\left(\widehat{\boldsymbol{\Theta}}(t)\right), \tag{3.79}$$

$$\dot{\widehat{\lambda}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}\left(E\left[L_i^\pi\right] - \widehat{\lambda}_i(t)\right), \tag{3.80}$$

$$\dot{\widehat{\sigma}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right)\right]}{\widehat{\lambda}_i(t)} - \widehat{\sigma}_i(t)\right), \tag{3.81}$$

$$\dot{\widehat{\rho}}_i(t) = \frac{1}{\tau_{slot}}\left(\frac{\overline{r}_i^*\left(\widehat{\mathbf{b}}(t)\right)}{\tau_{slot}} - \widehat{\rho}_i(t)\right), \tag{3.82}$$

where

$$u_i\left(\widehat{\mathbf{\Theta}}(t)\right)=\tau_{slot}\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t),F_i^\pi\right)\right)\right]}{E\left[r_i^*\left(\widehat{\mathbf{b}}(t),C^\pi\right)\right]}, \tag{3.83}$$

and $\mathbf{r}^*\left(\widehat{\mathbf{b}}(t),c\right)\in\mathcal{R}^*\left(\widehat{\mathbf{b}}(t),c\right)$ for each $c\in\mathcal{C}$. Here,

$$\left(\widehat{\mathbf{z}}^b\left(\widehat{\mathbf{\Theta}}(t)\right),\widehat{\mathbf{z}}^d\left(\widehat{\mathbf{\Theta}}(t)\right)\right)\in-\mathcal{Z}_{\mathcal{H}}\left(\mathbf{\Theta}\right). \tag{3.84}$$

Here $\widehat{z}_i^b\left(\widehat{\mathbf{\Theta}}(t)\right)$ and $\widehat{z}_i^d\left(\widehat{\mathbf{\Theta}}(t)\right)$ are terms mimicking the role of the operators $[.]_{\underline{b}}$ and $[.]_{\underline{d}}$ in (3.24) and (3.25), and ensure that $\left(\widehat{\mathbf{\Theta}}(t)\right)_{t\geq 0}$ stays in $\mathcal{H}$ (see Section 4.3 of [29] for a discussion about projected stochastic approximation). For $\mathbf{\Theta}=(\mathbf{m},\boldsymbol{\mu},\mathbf{v},\mathbf{b},\mathbf{d},\boldsymbol{\lambda},\boldsymbol{\sigma},\boldsymbol{\rho})\in\mathcal{H}$, $\mathcal{Z}_{\mathcal{H}}\left(\mathbf{\Theta}\right)\subset\mathbb{R}^{2N}$ is the set containing only the zero element when $(\mathbf{b},\mathbf{d})$ is in the interior of the set

$$\mathcal{H}_{\mathcal{BD}}=\left\{(\mathbf{b},\mathbf{d})\in\mathbb{R}^{2N}:\text{ for each } i\in\mathcal{N},\ b_i\geq\underline{b},\ d_i\geq\underline{d}\right\},$$

and for $(\mathbf{b},\mathbf{d})$ on the boundary of the set $\mathcal{H}_{\mathcal{BD}}$, $\mathcal{Z}_{\mathcal{H}}\left(\mathbf{\Theta}\right)$ is the convex cone generated by the outer normals at $(\mathbf{b},\mathbf{d})$ of the faces of $\mathcal{H}_{\mathcal{BD}}$ on which $(\mathbf{b},\mathbf{d})$ lies. Thus, for a given $\mathbf{\Theta}$, $-\mathcal{Z}_{\mathcal{H}}\left(\mathbf{\Theta}\right)$ contains reflection terms pointing in the right directions to keep $\left(\widehat{\mathbf{\Theta}}(t)\right)_{t\geq 0}$ in $\mathcal{H}$. For $\mathbf{\Theta}$ with $(\mathbf{b},\mathbf{d})$ in the interior of $\mathcal{H}_{\mathcal{BD}}$,

$$\widehat{z}_i^b=0,\ \widehat{z}_i^d=0,\ \forall\, i\in\mathcal{N},\ \forall\ \left(\widehat{\mathbf{z}}^b,\widehat{\mathbf{z}}^d\right)\in-\mathcal{Z}_{\mathcal{H}}\left(\mathbf{\Theta}\right), \tag{3.85}$$

Also note that for all $\mathbf{\Theta}\in\mathcal{H}$

$$\widehat{z}_i^b\geq 0,\ \widehat{z}_i^d\geq 0,\ \forall\, i\in\mathcal{N},\ \forall\ \left(\widehat{\mathbf{z}}^b,\widehat{\mathbf{z}}^d\right)\in-\mathcal{Z}_{\mathcal{H}}\left(\mathbf{\Theta}\right), \tag{3.86}$$

i.e., the components of all the terms in $-\mathcal{Z}_{\mathcal{H}}(\mathbf{\Theta})$ are non-negative and this is clear from the definition of the set $\mathcal{H}_{\mathcal{BD}}$ which indicates that the reflection terms are needed only needed when the parameters hit a lower bound. Thus, $\widehat{z}_i^b\left(\widehat{\mathbf{\Theta}}(t)\right)$ and $\widehat{z}_i^d\left(\widehat{\mathbf{\Theta}}(t)\right)$ are terms mimicking the role of the operators $[.]_{\underline{b}}$ and $[.]_{\underline{d}}$. Also, note that $u_i(.)$ is a set valued map (and hence (3.75)-(3.82) describes a differential inclusion) since the denominator $E\left[r_i^*\left(\widehat{\mathbf{b}}(t), C^\pi\right)\right]$ in (3.83) is a set valued map. Finally, note that the above definition only requires that $\left(\widehat{\mathbf{\Theta}}(t)\right)_{t\geq 0}$ is differentiable for *almost* all $t \geq 0$, i.e., we are considering the class of absolutely continuous functions $\left(\widehat{\mathbf{\Theta}}(t)\right)_{t\geq 0}$ that satisfy (3.75)-(3.82).

Although we will rigorously establish the relationship between the evolution of parameters of NOVA and (3.75)-(3.82) in the next subsection, we can see that the differential inclusion (3.75)-(3.82) reflects the average dynamics of the evolution of parameters in NOVA by comparing (3.75)-(3.82) against the update rules (3.21)-(3.26) and (3.72)-(3.73) in NOVA. For instance, this is apparent when we compare the update rule

$$\mu_{i,s_i+1} - \mu_{i,s_i} = \epsilon\left(\frac{l_{i,s_i+1}}{\lambda_{i,s_i}} q^*_{i,s_i+1} - \mu_{i,s_i}\right)$$

for NOVA parameter μ_{i,s_i+1} given in (3.22), against (3.76) describing the evolution of the parameter $\widehat{\mu}_i(t)$. Note that the rate of change of $\widehat{\mu}_i(t)$ given in (3.76) has a scaling term $\frac{1}{u_i\left(\widehat{\mathbf{\Theta}}(t)\right)}$ which corresponds to the segment download rate of video client i at time t (and $u_i\left(\widehat{\mathbf{\Theta}}(t)\right)$ defined in (3.83) corresponds to expected segment download time of video client i at time t). This scaling by segment download rate is naturally expected for the rate of change of parameters $\widehat{m}_i(t)$, $\widehat{\mu}_i(t)$, $\widehat{v}_i(t)$, $\widehat{d}_i(t)$, $\widehat{\lambda}_i(t)$, and $\widehat{\sigma}_i(t)$ which correspond to NOVA parameters of video client i that are updated when a segment download is completed, and thus we can view $\frac{1}{u_i\left(\widehat{\mathbf{\Theta}}(t)\right)}$ as the update rate associated with these parameters. Similarly, we can view the constant scaling term $\frac{1}{\tau_{slot}}$ in (3.82) describing the evolution of $\widehat{\rho}_i(t)$ as the corresponding update rate by noting that the associated (auxiliary) NOVA parameter $\rho_{i,k}$ is updated at the beginning of every slot, i.e., once every τ_{slot} seconds. Finally, note that (3.78) describing the evolution of $\widehat{b}_i(t)$ can be rewritten as

$$\dot{\widehat{b}}_i(t) = \frac{1}{\tau_{slot}} \left(\frac{\tau_{slot}}{(1+\overline{\beta}_i)} \right) - \frac{1}{u_i\left(\widehat{\mathbf{\Theta}}(t)\right)} \left(E\left[L_i^\pi\right]\right) + \widehat{z}_i^b\left(\widehat{\mathbf{\Theta}}(t)\right),$$

and presence of the two scaling terms $\frac{1}{\tau_{slot}}$ and $\frac{1}{u_i(\widehat{\mathbf{\Theta}}(t))}$ reflects the fact that the corresponding NOVA parameter $b_{i,k}$ is updated at the beginning of every slot (using (3.20)) and when a segment download of video client i is completed (using (3.24)). Thus, we can expect that (3.75)-(3.82) captures the average dynamics of NOVA, and the presence of the two video client dependent update rates $\frac{1}{\tau_{slot}}$ and $\frac{1}{u_i(\widehat{\mathbf{\Theta}}(t))}$ reflects the *asynchronous* nature of the evolution of parameters in NOVA where different video clients are updating their parameters at their own (possibly time varying) rates.

Now, we study the differential inclusion (3.75)-(3.82) to identify properties that will help us to study convergence behavior. The next result shows that the differential inclusion is 'well behaved'.

Lemma 3.6. *The differential inclusion* (3.75)-(3.82) *is well defined, i.e., there exists an absolutely continuous function that solves* (3.75)-(3.82) *for any* $\widehat{\mathbf{\Theta}}(0) \in \mathcal{H}$. *Further, these solutions are Lipschitz continuous and stay in* $\mathcal{H}$ *and hence are bounded.*

Proof. The existence of solution follows from the proof of Theorem 3.4 (in Subsection 3.5.3) where we obtain a solution satisfying (3.75)-(3.82). The boundedness follows from (3.84) and arguments similar to those in Lemma 3.1. The Lipschitz continuity of the paths follows from the fact that all the terms on the right hand side of (3.75)-(3.82) can be bounded above for $\widehat{\mathbf{\Theta}}(t) \in \mathcal{H}$.

□

Definition 3.1. *Stationary resource allocation policy:* Let $(\mathbf{r}(c))_{c\in\mathcal{C}}$ be a $|\mathcal{C}|$ length vector (of vectors) where $\mathbf{r}(c) \in \mathbb{R}_+^N$. We refer to $(\mathbf{r}(c))_{c\in\mathcal{C}}$ as a stationary resource allocation policy as we can associate $(\mathbf{r}(c))_{c\in\mathcal{C}}$ with a resource allocation policy that allocates resource $\mathbf{r}(c)$ in a slot k when $C_k = c$, and thus the policy carries out the resource allocation in a slot only based on the allocation constraint in the slot.

Definition 3.2. *Feasible stationary resource allocation policy:* We say that a stationary resource allocation policy $\left((\mathbf{r}(c))_{c\in\mathcal{C}}\right)$ is feasible if

$$\mathbf{r}(c) \geq \mathbf{r}_{\min} \text{ and } c(\mathbf{r}(c)) \leq 0,\ \forall\ c \in \mathcal{C}.$$

Definition 3.3. *Stationary quality adaptation policy for video client i:*
Let $(q_i(f_i))_{f_i\in\mathcal{F}_i} \in \mathbb{R}_+^{\mathcal{F}_i}$. We refer to $(q_i(f_i))_{f_i\in\mathcal{F}_i}$ as a stationary quality adaptation policy for video client $i \in \mathcal{N}$ as we can associate $(q_i(f_i))_{f_i\in\mathcal{F}_i}$ with a quality adaptation policy for video client i that chooses quality $q_i(f_i)$ for each segment s with QR tradeoff $F_{i,s} = f_i$, and thus the policy carries out quality adaptation for a segment based only on the QR tradeoff of that segment.

Definition 3.4. *Feasible stationary quality adaptation policy for video client i:* We say that a stationary quality adaptation policy $(q_i(f_i))_{f_i\in\mathcal{F}_i}$ for video client i is feasible if $0 \leq q_i(f_i) \leq q_{\max}$ for each $f_i \in \mathcal{F}_i$.

Next, we define the set $\widetilde{\mathcal{H}} \subset \mathbb{R}^{8N}$ as

$$\begin{aligned}\widetilde{\mathcal{H}} \ = \ \Bigg\{ & (\mathbf{m}, \boldsymbol{\mu}, \mathbf{v}, \mathbf{b}, \mathbf{d}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathcal{H} :\ \lambda_i = E[L_i^\pi]\ \forall\ i \in \mathcal{N}; \\ & \exists \text{ a feasible stationary resource allocation policy } (\mathbf{r}(c))_{c\in\mathcal{C}} \text{ s.t.} \\ & \frac{\mathbb{E}\left[r_i\left(C^\pi\right)\right]}{\tau_{slot}} = \rho_i\ \forall\ i \in \mathcal{N}; \\ & \text{for each } i \in \mathcal{N},\ \exists \text{ there is a feasible stationary quality} \\ & \text{adaptation scheme } \left((q_i(f_i))_{f_i\in\mathcal{F}_i}\right) \text{ such that } \frac{E\left[L_i^\pi q_i\left(F_i^\pi\right)\right]}{E\left[L_i^\pi\right]} = \mu_i, \\ & \mathrm{Var}\left(q_i\left(F_i^\pi\right)\right) \leq v_i \leq q_{\max}^2,\ \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right]}{\mathbb{E}\left[L_i^\pi\right]} \leq \sigma_i \leq f_{\max} \Bigg\}. \end{aligned} \tag{3.87}$$

We can view $\widetilde{\mathcal{H}}$ as the set of 'achievable' parameters in $\mathcal{H}$, i.e., for any element $(\mathbf{m}, \boldsymbol{\mu}, \mathbf{v}, \mathbf{b}, \mathbf{d}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathcal{H}$ there is some feasible stationary resource allocation policy with mean resource allocation per unit time $\boldsymbol{\rho}$, and there is some feasible stationary quality adaptation policy for each i that has a mean quality μ_i, variance in quality which is at least v_i and mean segment size which is at least σ_i (and satisfies $\lambda_i = E[L_i^\pi] \ \forall\ i \in \mathcal{N}$).

It can be verified that $\widetilde{\mathcal{H}}$ is a bounded, closed and convex set (using an approach similar to that in Lemma 5 (b) in [26]). Hence, we conclude that for any $\boldsymbol{\Theta} \in \mathcal{H}$, there exists a unique projection of $\widetilde{\boldsymbol{\Theta}} \in \mathcal{H}$ on the set $\widetilde{\mathcal{H}}$. Let $\widetilde{.}$ denote this projection operator. Hence, for any $\boldsymbol{\Theta} \in \mathcal{H}$, $d_{8N}\left(\boldsymbol{\Theta}, \widetilde{\mathcal{H}}\right) = d_{8N}\left(\boldsymbol{\Theta}, \widetilde{\boldsymbol{\Theta}}\right)$. The next result states that, irrespective of the initialization, the differential inclusion converges to the bounded, closed and convex set $\widetilde{\mathcal{H}}$ of achievable parameters.

Lemma 3.7. *There exist finite constants $\chi_0 > 0$ and χ_1 such that for any initialization $\widehat{\boldsymbol{\Theta}}(0) \in \mathcal{H}$,*

$$\frac{d}{dt} d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \mathcal{H}\right) \leq -\chi_0 d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \mathcal{H}\right) + \chi_1 d_N\left(\widehat{\boldsymbol{\lambda}}(t), \boldsymbol{\lambda}^\pi\right).$$

Hence,

$$\lim_{t\to\infty} d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \mathcal{H}\right) = 0.$$

Proof. The result is an application of a generalization of Lemma 3 in [49]. It can be proved by modifying the arguments in [49], and using the update rule (3.80) for $\widehat{\boldsymbol{\lambda}}(t)$ which can be also written as

$$\dot{\widehat{\lambda}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} d_1\left(\widehat{\lambda}_i(t), E\left[L_i^\pi\right]\right).$$

□

In the next result, we provide the main convergence result for the differential inclusion (3.75)-(3.82) which states that $\widehat{\boldsymbol{\Theta}}(t)$ converges to the following set

$$\begin{aligned}\mathcal{H}^* &= \Big\{(\mathbf{m},\boldsymbol{\mu},\mathbf{v},\mathbf{b},\mathbf{d},\boldsymbol{\lambda},\boldsymbol{\sigma},\boldsymbol{\rho}) \in \mathcal{H} : \Big(\boldsymbol{\rho}, \left(h_i^B\left(b_i\right)\right)_{i\in\mathcal{N}}, \left(h_i^D\left(d_i\right)\right)_{i\in\mathcal{N}}\Big) \in \mathcal{X}^\pi, \\ &\quad \text{and for each } i \in \mathcal{N},\ m_i = \mu_i = m_i^\pi,\ v_i = v_i^\pi\} \end{aligned} \tag{3.88}$$

Recall that Theorem 3.2 suggested that we can prove Theorem 3.1, if we can show that the updates (3.21)-(3.26) guide NOVA parameters $(\boldsymbol{\theta}_{i,s})_{s\geq 1}$ of video client i to the set $\mathcal{H}_i^*$ (defined in (3.70)) for each video client $i \in \mathcal{N}$. Note that for each $i \in \mathcal{N}$, $\mathcal{H}_i^*$ is a set obtained by projecting $\mathcal{H}^*$ on a lower dimensional space (by considering only video client i's components and 'dropping' the components $(\boldsymbol{\lambda},\boldsymbol{\sigma},\boldsymbol{\rho})$). Hence, the following result along with Theorem 3.4 (which relates evolution of NOVA parameters to the differential inclusion) help us to establish the desired convergence property for NOVA parameters.

Theorem 3.3. *(a) For* $\widehat{\boldsymbol{\Theta}} = \left(\widehat{\mathbf{m}},\widehat{\boldsymbol{\mu}},\widehat{\mathbf{v}},\widehat{\mathbf{b}},\widehat{\mathbf{d}},\widehat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\sigma}},\widehat{\boldsymbol{\rho}}\right) \in \mathcal{H}$*, and some* $(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi, \mathbf{d}^\pi) \in \mathcal{X}^\pi$*, let*

$$\begin{aligned} L\left(\widehat{\boldsymbol{\Theta}}\right) &:= -\sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i U_i^E\left(\widehat{\mu}_i - U_i^V\left(\widehat{v}_i\right)\right) + \sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i\left(\widehat{m}_i - m_i^\pi\right)^2 \\ &\quad + \sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i d_i^\pi\left(\frac{p_i^d\widehat{\sigma}_i}{\overline{p}_i} - 1\right) + \sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\int_{\underline{d}}^{\widehat{d}_i}\left(h_i^D(e) - d_i^\pi\right)de \\ &\quad + \sum_{i\in\mathcal{N}}\left(\widehat{\lambda}_i b_i^\pi\widehat{\sigma}_i - \tau_{slot} b_i^\pi\widehat{\rho}_i\right) + \sum_{i\in\mathcal{N}}\sigma_i^\pi\int_{\underline{b}}^{\widehat{b}_i}\left(h_i^B(e) - b_i^\pi\right)de \\ &\quad + \frac{\chi_2}{\chi_0} d_{8N}\left(\widehat{\boldsymbol{\Theta}}, \widetilde{\mathcal{H}}\right), \end{aligned} \tag{3.89}$$

where χ_0 *is the positive constant from Lemma 3.7, and* χ_2 *is a large positive constant* *(the value is given in the proof). If* $\widehat{\boldsymbol{\Theta}}(0) \in \mathcal{H}$*, then for almost all* t

$$\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} \begin{cases} \leq 0, \ \forall \ \widehat{\boldsymbol{\Theta}}(t) \in \mathcal{H}, \\ < 0, \ \forall \ \widehat{\boldsymbol{\Theta}}(t) \notin \mathcal{H}^*. \end{cases}$$

(b) If $\widehat{\boldsymbol{\Theta}}(0) \in \mathcal{H}$*, then*

$$\lim_{t\to\infty} d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \mathcal{H}^*\right) = 0.$$

Proof. The proof of part (b) relies on the analysis of the drift of the Lyapunov function $L(.)$ defined in (3.89) of part (a) for $\widehat{\boldsymbol{\Theta}} = \left(\widehat{\mathbf{m}}, \widehat{\boldsymbol{\mu}}, \widehat{\mathbf{v}}, \widehat{\mathbf{b}}, \widehat{\mathbf{d}}, \widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{\sigma}}, \widehat{\boldsymbol{\rho}}\right) \in \mathcal{H}$ where $(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi, \mathbf{d}^\pi) \in \mathcal{X}^\pi$ (defined in (3.69)). Here χ_0 is the positive constant from Lemma 3.7, and χ_2 is a positive constant whose value is chosen to be large enough to satisfy certain conditions, and is specified towards the end of the proof (above (3.101)). The choice of several terms in the Lyapunov function $L(.)$ given above are motivated by the choice of Lyapunov functions in [24] and [50]. The first term $-\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i U_i^E\left(\widehat{\mu}_i - U_i^V\left(\widehat{v}_i\right)\right)$ is similar to terms in the Lyapunov function in [24], and resembles an estimate for a scaled version of the objective of OPT(S) (see (3.7)) since $\widehat{\mu}_i(.)$ tracks mean quality and $\widehat{v}_i(.)$ tracks variance in quality so that a negative drift would suggest that the estimate is decreasing and we are moving in the right direction. The terms in the second line $\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i d_i^\pi\left(\frac{p_i^d\widehat{\sigma}_i}{\overline{p}_i} - 1\right) + \sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\int_{\underline{d}}^{\widehat{d}_i}\left(h_i^D(e) - d_i^\pi\right)de$ are similar to the terms chosen in the Lyapunov function in [50]. However, note that our choice and analysis of the Lyapunov function has many novel elements. For instance, the term $\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i\left(\widehat{m}_i - m_i^\pi\right)^2$ (which can be viewed as weighted distance of '$\widehat{\mathbf{m}}$' component in $\widehat{\boldsymbol{\Theta}}(t)$ to '$\widehat{\mathbf{m}}$' component of elements in $\mathcal{H}^*$) allows us to accommodate objectives involving variability terms (i.e., non-zero $U_i^V\left(\widehat{v}_i\right)$) and the terms in $\sum_{i\in\mathcal{N}} b_i^\pi\left(\widehat{\lambda}_i\widehat{\sigma}_i - \tau_{slot}\widehat{\rho}_i\right) + \sum_{i\in\mathcal{N}}\sigma_i^\pi\int_{\underline{b}}^{\widehat{b}_i}\left(h_i^B(e) - b_i^\pi\right)de$ which allow us to accommodate the rebuffering constraints (which involves comparing averages over time time scales). Further, note that our analysis of $L(.)$ will establish a convergence result for a differential inclusion associated with an algorithm NOVA which, unlike those in [24] and [50], uses *asynchronous* updates.

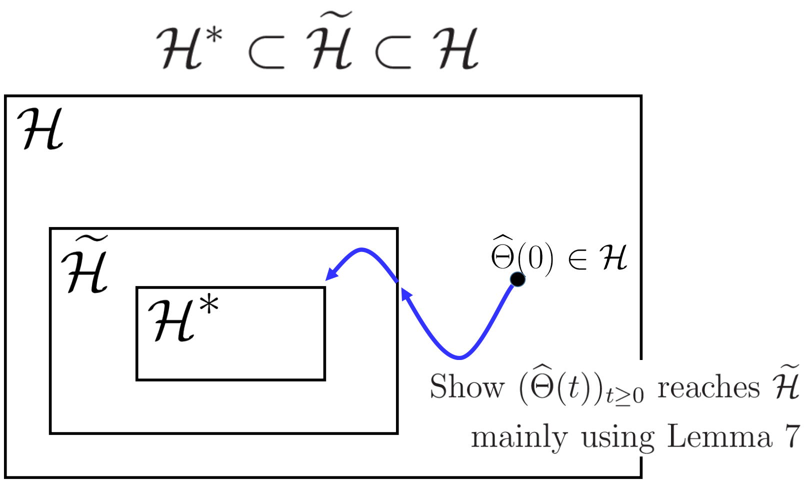

Figure 3.3: $\mathcal{H}^* \subset \widetilde{\mathcal{H}} \subset \mathcal{H}$, and we show that $\widehat{\boldsymbol{\Theta}}(t))_{t\geq 0}$ reaches $\widetilde{\mathcal{H}}$ mainly using Lemma 3.7

In the first part of the proof, we establish that the Lyapunov function L has a non-positive drift, and that the drift is strictly negative outside $\mathcal{H}^*$. Note that since $(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi, \mathbf{d}^\pi) \in \mathcal{X}^\pi$, there is some optimal solution $\left(\left((q_i^\pi(f))_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, (\mathbf{r}^\pi(c))_{c\in\mathcal{C}}\right)$ to OPTSTAT with $\rho_i^\pi = \mathbb{E}\left[r_i^\pi(C^\pi)\right]$ for each $i \in \mathcal{N}$ and $\mathbf{b}^\pi$ and $\mathbf{d}^\pi$ as the associated optimal Lagrange multipliers for the constraints (3.43) and (3.44) respectively.

The proof is a bit lengthy, and the following is rough outline of the initial steps in the proof (RHS is shorthand for right hand side):

$$\begin{aligned}\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} &\leq \text{ RHS of (3.90)} \\ &\leq \text{ RHS of (3.91)} \\ &\leq \text{ RHS of (3.95)} \\ &\leq \text{ RHS of (3.100)} \\ &\leq \text{ RHS of (3.102)}\end{aligned}$$

We use definition of $L(.)$ and (3.75)-(3.82) to obtain (3.90). We obtain (3.91) mainly counting on (optimality) properties of optimal solution to $\text{QNOVA}_i\left(\widehat{\boldsymbol{\theta}}_i(t), f_i\right)$ and Lemma 3.3 (c), and we obtain (3.95) mainly counting on (optimality) properties of optimal solutions to $\text{RNOVA}\left(\widehat{\mathbf{b}}(t), c\right)$ and $\text{QNOVA}_i\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$, where $\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t) = (m_i^\pi, \widehat{\mu}_i(t), \widehat{v}_i(t), \widehat{b}_i(t), \widehat{d}_i(t))$. In step (3.100), we mainly collect projection (projection of $\widehat{\boldsymbol{\Theta}}(t)$ on $\widetilde{\mathcal{H}}$) error terms and terms containing $\left|\widehat{\lambda}_i(t) - \lambda_i^\pi\right|$, bound them and nullify the role of these terms by using (3.80) and Lemma 3.7 and picking large enough χ_2. Finally, we obtain (3.102) mainly counting on properties of optimal solution and optimal Lagrange multipliers of OPTSTAT. In (3.102), we have an upper bound for $\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt}$ as a sum of several functions which are shown to be non-positive using aforementioned properties, and we use additional arguments to establish strict negativity of $\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt}$ outside $\mathcal{H}^*$ and conclude proof of part (a).

Using the definition of $L(.)$ and (3.75)-(3.82), we have that

$$\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} \leq -\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(U_i^E\right)'\left(\widehat{\mu}_i(t)-U_i^V\left(\widehat{v}_i(t)\right)\right) \tag{3.90}$$

$$\left(\frac{1}{u_i(t)}\left(\frac{E\left[L_i^\pi q_i^*(t)\right]}{\widehat{\lambda}_i(t)}-\widehat{\mu}_i(t)\right)\right.$$

$$\left.-\frac{\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)}{u_i(t)}\left(\frac{E\left[L_i^\pi\left(q_i^*(t)-\widehat{m}_i(t)\right)^2\right]}{\widehat{\lambda}_i(t)}-\widehat{v}_i(t)\right)\right)$$

$$+\sum_{i\in\mathcal{N}}2\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\frac{\left(U_i^E\right)'\left(\widehat{\mu}_i(t)-U_i^V\left(\widehat{v}_i(t)\right)\right)\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)}{u_i(t)}$$

$$\left(\widehat{m}_i(t)-m_i^\pi\right)\left(\frac{E\left[L_i^\pi q_i^*(t)\right]}{\widehat{\lambda}_i(t)}-\widehat{m}_i(t)\right)$$

$$+\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\frac{1}{u_i(t)}\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*(t)\right)\right]}{\widehat{\lambda}_i(t)}-\widehat{\sigma}_i(t)\right)$$

$$+\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right)-d_i^\pi\right)\frac{1}{u_i(t)}\left(p_i^d\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*(t)\right)\right]}{\widehat{\lambda}_i(t)\overline{p}_i}-1\right)$$

$$+\sum_{i\in\mathcal{N}}\widehat{\lambda}_i(t)b_i^\pi\frac{1}{u_i(t)}\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*(t)\right)\right]}{\widehat{\lambda}_i(t)}-\widehat{\sigma}_i(t)\right)-\sum_{i\in\mathcal{N}}b_i^\pi\left(\frac{E\left[r_i^*(t)\right]}{\tau_{slot}}-\widehat{\rho}_i(t)\right)$$

$$+\sum_{i\in\mathcal{N}}\sigma_i^\pi\left(h_i^B\left(\widehat{b}_i(t)\right)-b_i^\pi\right)\left(\frac{1}{\left(1+\overline{\beta}_i\right)}-\frac{E\left[L_i^\pi\right]}{u_i(t)}\right)$$

$$-\chi_2 d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t),\widetilde{\mathcal{H}}\right)+\frac{\chi_2\chi_1}{\chi_0}d_N\left(\widehat{\boldsymbol{\lambda}}(t),\boldsymbol{\lambda}^\pi\right)+\sum_{i\in\mathcal{N}}\dot{\widehat{\lambda}}_i(t)\widetilde{l}_{i1}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$$

$$-+\sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right)-b_i^\pi\right)\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*(t)\right)\right]}{u_i(t)}-\frac{E\left[r_i^*(t)\right]}{\tau_{slot}}\right),$$

where we have collected terms involving $\dot{\widehat{\lambda}}_i(t)$ and grouped them together in the term $\dot{\widehat{\lambda}}_i(t)\widetilde{l}_{i1}\left(\widehat{\mathbf{\Theta}}(t)\right)$ so that for each $i \in \mathcal{N}$,

$$\begin{aligned}\widetilde{l}_{1i}\left(\widehat{\mathbf{\Theta}}\right) &= -\left(1+\overline{\beta}_i\right) U_i^E\left(\widehat{\mu}_i - U_i^V\left(\widehat{v}_i\right)\right) + \left(1+\overline{\beta}_i\right)\left(\widehat{m}_i - m_i^\pi\right)^2 \\ &\quad + \sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right) d_i^\pi \left(\frac{p_i^d \widehat{\sigma}_i}{\overline{p}_i} - 1\right) + b_i^\pi \widehat{\sigma}_i.\end{aligned}$$

For brevity, we have not explicitly indicated the dependence of many terms above on $\widehat{\mathbf{\Theta}}(t)$. For instance, $u_i\left(\widehat{\mathbf{\Theta}}(t)\right)$ is shorthand for $u_i(t)$, and $E\left[L_i^\pi q_i^*(t)\right]$ is shorthand for $E\left[L_i^\pi q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right]$ where $\widehat{\boldsymbol{\theta}}_i(t) = (\widehat{m}_i(t), \widehat{\mu}_i(t), \widehat{v}_i(t), \widehat{b}_i(t), \widehat{d}_i(t))$ for each $i \in \mathcal{N}$. Also, note that we also added

$$\sum_{i\in\mathcal{N}} \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^*(t)\right)\right]}{u_i(t)} - \frac{E\left[r_i^*(t)\right]}{\tau_{slot}}\right)$$

to the right hand side of (3.90) where $E\left[r_i^*(t)\right]$ is a shorthand for $E\left[r_i^*\left(\widehat{\mathbf{b}}(t), C^\pi\right)\right]$ where $\mathbf{r}^*\left(\widehat{\mathbf{b}}(t), c\right) \in \mathcal{R}^*\left(\widehat{\mathbf{b}}(t), c\right)$ which is $\overline{\mathcal{R}}^*\left(\widehat{\mathbf{b}}(t)\right)$ for each $c \in \mathcal{C}$. This does not change the inequality since this expression evaluates to zero from the definition of $u_i(.)$ in (3.83). We also dropped the terms $\sum_{i\in\mathcal{N}} \sigma_i^\pi \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right) \widehat{z}_i^b\left(\widehat{\mathbf{\Theta}}(t)\right)$ and $\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right) - d_i^\pi\right)\widehat{z}_i^d\left(\widehat{\mathbf{\Theta}}(t)\right)$ from right hand side of (3.90) noting that they are less than or equal to zero. To see this, note that $\widehat{z}_i^b\left(\widehat{\mathbf{\Theta}}(t)\right) \geq 0$ (from (3.86)) which is equal to zero unless $\widehat{b}_i(t) = \underline{b}$ (from (3.85)) for which $h_i^B\left(\widehat{b}_i(t)\right) = 0 \leq b_i^\pi$. Hence, $\sum_{i\in\mathcal{N}} \sigma_i^\pi \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\widehat{z}_i^b\left(\widehat{\mathbf{\Theta}}(t)\right) \leq 0$, and we can similarly show that $\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right) - d_i^\pi\right)\widehat{z}_i^d\left(\widehat{\mathbf{\Theta}}(t)\right) \leq 0$.

Consider the right hand side of (3.90), and group the terms containing $q_i^*(t)$ except those in the fifth line (i.e., the line with the term $(\widehat{m}_i(t) - m_i^\pi)$) to note that we have negative of a scaled (by $\left(1+\overline{\beta}_i\right)/u_i(t)$) version of the expectation of the objective of QNOVA$_i\left(\widehat{\boldsymbol{\theta}}_i(t), f_i\right)$, i.e., $E\left[L_i^\pi \phi^Q\left(q_i^*(t), \widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right]$. Recall that $q_i^*(t)$ in the above calculations is a shorthand for $q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)$. Now, let $q_i^{*,m_i^\pi}(t)$ denote the shorthand for $q_i^*\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), F_i^\pi\right)$ where $\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t) = (m_i^\pi, \widehat{\mu}_i(t), \widehat{v}_i(t), \widehat{b}_i(t), \widehat{d}_i(t))$, i.e., $\widehat{\boldsymbol{\theta}}_i(t)$ with the first component set to m_i^π (defined in (3.65)). Next, we replace $q_i^*(t)$ appearing in the above inequality with $q_i^{*,m_i^\pi}(t)$, incorporate the correction term associated with this replacement into a function $\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right)$, and rewrite (3.90) as

$$
\begin{aligned}
\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} \leq\ & \Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right) - \sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(U_i^E\right)'\left(\widehat{\mu}_i(t) - U_i^V\left(\widehat{v}_i(t)\right)\right) \\
& \left(\frac{1}{u_i(t)}\left(\frac{E\left[L_i^\pi q_i^{*,m_i^\pi}(t)\right]}{\widehat{\lambda}_i(t)} - \widehat{\mu}_i(t)\right)\right. \\
& \left. -\frac{\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)}{u_i(t)}\left(\frac{E\left[L_i^\pi\left(q_i^{*,m_i^\pi}(t) - m_i^\pi\right)^2\right]}{\widehat{\lambda}_i(t)} - \widehat{v}_i(t)\right)\right) \\
& + \sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t) d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\frac{1}{u_i(t)}\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^{*,m_i^\pi}(t)\right)\right]}{\widehat{\lambda}_i(t)} - \widehat{\sigma}_i(t)\right) \\
& + \sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right) - d_i^\pi\right)\frac{1}{u_i(t)}\left(p_i^d\frac{E\left[L_i^\pi F_i^\pi\left(q_i^{*,m_i^\pi}(t)\right)\right]}{\widehat{\lambda}_i(t)\overline{p}_i} - 1\right) \\
& + \sum_{i\in\mathcal{N}}\widehat{\lambda}_i(t) b_i^\pi\frac{1}{u_i(t)}\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^{*,m_i^\pi}(t)\right)\right]}{\widehat{\lambda}_i(t)} - \widehat{\sigma}_i(t)\right) \\
& - \sum_{i\in\mathcal{N}} b_i^\pi\left(\frac{E\left[r_i^*(t)\right]}{\tau_{slot}} - \widehat{\rho}_i(t)\right) \\
& + \sum_{i\in\mathcal{N}}\sigma_i^\pi\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{1}{\left(1+\overline{\beta}_i\right)} - \frac{E\left[L_i^\pi\right]}{u_i(t)}\right) \\
& + \sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{E\left[L_i^\pi F_i^\pi\left(q_i^{*,m_i^\pi}(t)\right)\right]}{u_i(t)} - \frac{E\left[r_i^*(t)\right]}{\tau_{slot}}\right) \\
& - \chi_2 d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \widetilde{\mathcal{H}}\right) + \frac{\chi_2\chi_1}{\chi_0} d_N\left(\widehat{\boldsymbol{\lambda}}(t), \boldsymbol{\lambda}^\pi\right) + \sum_{i\in\mathcal{N}}\dot{\widehat{\lambda}}_i(t) l_{1i}\left(\widehat{\boldsymbol{\Theta}}(t)\right),
\end{aligned} \tag{3.91}
$$

where

$$\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right) = -\sum_{i\in\mathcal{N}} \frac{(1+\overline{\beta}_i)}{u_i(t)} E\left[L_i^\pi\left(\phi^Q\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right), \widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right.\right. \tag{3.92}$$
$$-\phi^Q\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), F_i^\pi\right), \widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), F_i^\pi\right)$$
$$\left.\left. -\left(U_i^E\right)'\left(\widehat{\mu}_i(t) - U_i^V\left(\widehat{v}_i(t)\right)\right)\left(U_i^V\right)'\left(\widehat{v}_i(t)\right) 2\left(\widehat{m}_i(t) - m_i^\pi\right)\left(q_i^*(t) - \widehat{m}_i(t)\right)\right)\right]$$

and for each $i \in \mathcal{N}$,

$$\begin{aligned} l_{1i}\left(\widehat{\boldsymbol{\Theta}}(t)\right) &= \tilde{l}_{1i}\left(\widehat{\boldsymbol{\Theta}}(t)\right) \\ &+2\frac{(1+\overline{\beta}_i)}{u_i(t)}\left(U_i^E\right)'\left(\widehat{\mu}_i(t) - U_i^V\left(\widehat{v}_i(t)\right)\right)\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)\left(\widehat{m}_i(t) - m_i^\pi\right)\widehat{m}_i(t). \end{aligned}$$

Note that we have included the terms in the third line of (3.90) in (3.92) after replacing $\widehat{\lambda}_i(t)$ with $E\left[L_i^\pi\right]$, and the correction for the modification is included as the second term of $l_{1i}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ defined above. From the definition (3.83) of $u_i(t)$, we have that

$$u_{\min} := \frac{\tau_{slot} l_{\min} f_{\min}}{r_{\max}}, \quad u_{\max} := \frac{\tau_{slot} l_{\max} f_{\max}}{r_{\min}} \tag{3.93}$$

are lower and upper bounds respectively on $u_i(t)$ for each $i \in \mathcal{N}$. Note that, in the term $\sum_{i\in\mathcal{N}} \dot{\widehat{\lambda}}_i(t) l_{1i}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$, we are collecting the terms scaled by $\dot{\widehat{\lambda}}_i(t)$ together so as to bound it by picking a large enough χ_2 since $\dot{\widehat{\lambda}}_i(t) \leq \frac{1}{u_{\min}} d_N\left(\widehat{\boldsymbol{\lambda}}(t), \boldsymbol{\lambda}^\pi\right) \leq \frac{1}{u_{\min}} d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \widetilde{\mathcal{H}}\right)$. Using the bounded nature of the terms involved in $l_{1i}\left(\widehat{\boldsymbol{\Theta}}\right)$ and $\tilde{l}_{1i}\left(\widehat{\boldsymbol{\Theta}}\right)$, we can show that there is some finite χ_3 such that

$$\frac{\chi_2\chi_1}{\chi_0} d_N\left(\widehat{\boldsymbol{\lambda}}(t), \boldsymbol{\lambda}^\pi\right) + \sum_{i\in\mathcal{N}} l_{1i}\left(\widehat{\boldsymbol{\Theta}}\right)\dot{\widehat{\lambda}}_i(t) \leq \chi_3 d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \widetilde{\mathcal{H}}\right) \tag{3.94}$$

holds for any $\widehat{\boldsymbol{\Theta}}(t) \in \mathcal{H}$.

If we group the terms containing $q_i^{*,m_i^\pi}(t)$ and $\mathbf{r}^*(t)$, we find that the right hand side of (3.91) contains negative of scaled versions of optimal value of objective functions of $\text{QNOVA}_i\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$ (i.e., $\phi^Q\left(q_i^*(t), \widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$) and those of $\text{RNOVA}\left(\widehat{\mathbf{b}}(t), c\right)$ (i.e., $\phi^R\left(\mathbf{r}^*(t), \widehat{\mathbf{b}}(t), c\right)$). Now using the optimality of $q_i^{*,m_i^\pi}(t)$ and $\mathbf{r}^*(t)$ with respect to $\text{QNOVA}_i\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$ and $\text{RNOVA}\left(\widehat{\mathbf{b}}(t), c\right)$, and using the fact that $q_i^\pi(f)$ and $\mathbf{r}^\pi(c)$ are feasible solutions for these optimization problems, we obtain the following inequality from (3.91) (obtained by replacing $q_i^{*,m_i^\pi}(t)$ and $\mathbf{r}^*(t)$ with $q_i^\pi(f)$ and $\mathbf{r}^\pi(c)$ in (3.91) and adding the correction term $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ associated with this replacement)

$$
\begin{aligned}
\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} &\leq \Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right) + \Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) \qquad (3.95)\\
&- \sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(U_i^E\right)'\left(\widetilde{\widehat{\mu}_i(t)} - U_i^V\left(\widetilde{\widehat{v}_i(t)}\right)\right)\left(\frac{1}{u_i(t)}\left(\frac{m_i^\pi E\left[L_i^\pi\right]}{\widehat{\lambda}_i(t)} - \widetilde{\widehat{\mu}_i(t)}\right)\right.\\
&\qquad\left. - \frac{\left(U_i^V\right)'\left(\widetilde{\widehat{v}_i(t)}\right)}{u_i(t)}\left(\frac{E\left[L_i^\pi\left(q_i^\pi\left(F_i^\pi\right) - m_i^\pi\right)^2\right]}{\widehat{\lambda}_i(t)} - \widetilde{\widehat{v}_i(t)}\right)\right)\\
&+ \sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t) d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\frac{1}{u_i(t)}\left(\frac{\sigma_i^\pi E\left[L_i^\pi\right]}{\widehat{\lambda}_i(t)} - \widetilde{\widehat{\sigma}_i(t)}\right)\\
&+ \sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right) - d_i^\pi\right)\frac{1}{u_i(t)}\left(p_i^d\frac{\sigma_i^\pi E\left[L_i^\pi\right]}{\widehat{\lambda}_i(t)\overline{p}_i} - 1\right)\\
&+ \sum_{i\in\mathcal{N}} \widehat{\lambda}_i(t) b_i^\pi \frac{1}{u_i(t)}\left(\frac{\sigma_i^\pi E\left[L_i^\pi\right]}{\widehat{\lambda}_i(t)} - \widetilde{\widehat{\sigma}_i(t)}\right) - \sum_{i\in\mathcal{N}} b_i^\pi\left(\frac{\rho_i^\pi}{\tau_{slot}} - \widetilde{\widehat{\rho}_i(t)}\right)\\
&+ \sum_{i\in\mathcal{N}} \sigma_i^\pi\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{1}{\left(1+\overline{\beta}_i\right)} - \frac{E\left[L_i^\pi\right]}{u_i(t)}\right)\\
&+ \sum_{i\in\mathcal{N}} \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{\sigma_i^\pi E\left[L_i^\pi\right]}{u_i(t)} - \frac{\rho_i^\pi}{\tau_{slot}}\right)\\
&+ + \frac{\chi_2\chi_1}{\chi_0} d_N\left(\widehat{\boldsymbol{\lambda}}(t), \boldsymbol{\lambda}^\pi\right) + \sum_{i\in\mathcal{N}} \dot{\widehat{\lambda}}_i(t) l_{1i}\left(\widehat{\boldsymbol{\Theta}}(t)\right) - \chi_2 d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \widetilde{\mathcal{H}}\right) + l_2\left(\widehat{\boldsymbol{\Theta}}(t)\right),
\end{aligned}
$$

where m_i^π, v_i^π and σ_i^π are defined in (3.65)-(3.67), (and $\boldsymbol{\rho}^\pi$ was chosen at the beginning of the proof- see below (3.89))

$$\begin{aligned}\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) &= -\frac{1}{\tau_{slot}}E\left[\phi^R\left(\mathbf{r}^*(t),\widehat{\mathbf{b}}(t),C^\pi\right)-\phi^R\left(\mathbf{r}^\pi\left(C^\pi\right),\widehat{\mathbf{b}}(t),C^\pi\right)\right] \\ &\quad -\sum_{i\in\mathcal{N}}\frac{\left(1+\overline{\beta}_i\right)}{u_i(t)}E\left[L_i^\pi\left(\phi^Q\left(q_i^*\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t),F_i^\pi\right),\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t),F_i^\pi\right)\right.\right. \\ &\qquad \left.\left.-\phi^Q\left(q_i^\pi\left(F_i^\pi\right),\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t),F_i^\pi\right)\right)\right],\end{aligned} \tag{3.96}$$

and $\widetilde{\cdot}$ is the projection operator mapping elements in $\mathcal{H}$ to the set $\widetilde{\mathcal{H}}$. Here

$$\left(\widetilde{\widehat{\mathbf{m}}(t)},\widetilde{\widehat{\boldsymbol{\mu}}(t)},\widetilde{\widehat{\mathbf{v}}(t)},\widetilde{\widehat{\mathbf{b}}(t)},\widetilde{\widehat{\mathbf{d}}(t)},\widetilde{\widehat{\boldsymbol{\lambda}}(t)},\widetilde{\widehat{\boldsymbol{\sigma}}(t)},\widetilde{\widehat{\boldsymbol{\rho}}(t)}\right) := \widetilde{\widehat{\boldsymbol{\Theta}}(t)}.$$

Due to the definition of $\widetilde{\mathcal{H}}$ (see (3.87)), $\widetilde{\widehat{\mathbf{m}}(t)}=\widehat{\mathbf{m}}(t)$, $\widetilde{\widehat{\mathbf{b}}(t)}=\widehat{\mathbf{b}}(t)$ and $\widetilde{\widehat{\mathbf{d}}(t)}=\widehat{\mathbf{d}}(t)$ as (3.87) does not impose any additional restrictions on these components. Also, for each $i\in\mathcal{N}$,

$$\begin{aligned}\widetilde{\widehat{\boldsymbol{\theta}}_i(t)} &:= \left(\widehat{m}_i(t),\widetilde{\widehat{\mu}_i(t)},\widetilde{\widehat{v}_i(t)},\widehat{b}_i(t),\widehat{d}_i(t)\right), \\ \widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t) &:= \left(m_i^\pi,\widetilde{\widehat{\mu}_i(t)},\widetilde{\widehat{v}_i(t)},\widehat{b}_i(t),\widehat{d}_i(t)\right).\end{aligned}$$

Note that we have replaced components of $\widehat{\boldsymbol{\mu}}(t)$, $\widehat{\mathbf{v}}(t)$, $\widehat{\boldsymbol{\lambda}}(t)$, $\widehat{\boldsymbol{\sigma}}(t)$ and $\widehat{\boldsymbol{\rho}}(t)$ appearing in (3.91) with those of $\widetilde{\widehat{\boldsymbol{\mu}}(t)}$, $\widetilde{\widehat{\mathbf{v}}(t)}$, $\widetilde{\widehat{\boldsymbol{\lambda}}(t)}$, $\widetilde{\widehat{\boldsymbol{\sigma}}(t)}$ and $\widetilde{\widehat{\boldsymbol{\rho}}(t)}$ respectively, and in (3.95), we have added the function $l_2(.)$ defined below to account for these replacements:

$$
\begin{aligned}
l_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) = & -\sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(\left(U_i^E\right)'\left(\widehat{\mu}_i(t)-U_i^V\left(\widehat{v}_i(t)\right)\right)\right. \\
& \qquad -\left(U_i^E\right)'\left(\widetilde{\widehat{\mu}_i(t)}-U_i^V\left(\widetilde{\widehat{v}_i(t)}\right)\right)\Bigg) \\
& \qquad \left(\frac{1}{u_i(t)}\left(\frac{m_i^\pi E\left[L_i^\pi\right]}{\widehat{\lambda}_i(t)}-\widetilde{\widehat{\mu}_i(t)}\right)\right. \\
& \qquad \left.-\frac{\left(U_i^V\right)'\left(\widetilde{\widehat{v}_i(t)}\right)}{u_i(t)}\left(\frac{E\left[L_i^\pi\left(q_i^\pi\left(F_i^\pi\right)-m_i^\pi\right)^2\right]}{\widehat{\lambda}_i(t)}-\widetilde{\widehat{v}_i(t)}\right)\right) \\
& -\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(U_i^E\right)'\left(\widehat{\mu}_i(t)-U_i^V\left(\widehat{v}_i(t)\right)\right) \\
& \left(\frac{\left(U_i^V\right)'\left(\widetilde{\widehat{v}_i(t)}\right)-\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)}{u_i(t)}\left(\frac{E\left[L_i^\pi\left(q_i^\pi\left(F_i^\pi\right)-m_i^\pi\right)^2\right]}{\widehat{\lambda}_i(t)}-\widetilde{\widehat{v}_i(t)}\right)\right) \\
& -\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(U_i^E\right)'\left(\widehat{\mu}_i(t)-U_i^V\left(\widehat{v}_i(t)\right)\right) \\
& \left(\left(\widetilde{\widehat{\mu}_i(t)}-\widehat{\mu}_i(t)\right)-\frac{\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)}{u_i(t)}\left(\widetilde{\widehat{v}_i(t)}-\widehat{v}_i(t)\right)\right) \\
& +\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\frac{1}{u_i(t)}\left(\widetilde{\widehat{\sigma}_i(t)}-\widehat{\sigma}_i(t)\right) \\
& -\sum_{i\in\mathcal{N}}\widehat{\lambda}_i(t)b_i^\pi\frac{1}{u_i(t)}\left(\widetilde{\widehat{\sigma}_i(t)}-\widehat{\sigma}_i(t)\right)-\sum_{i\in\mathcal{N}}b_i^\pi\left(\widetilde{\widehat{\rho}_i(t)}-\widehat{\rho}_i(t)\right).
\end{aligned} \tag{3.97}
$$

Using the bounded nature of the terms involved in the above expression for $l_2\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ (note that here we are using the fact that the functions $\left(U_i^V, U_i^E\right)_{i\in\mathcal{N}}$ have Lipschitz continuous derivatives due to Assumptions U.V and U.E), we can show that there exists some large enough finite constant χ_4 such that

$$
l_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) \leq \chi_4 d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \widetilde{\mathcal{H}}\right) \tag{3.98}
$$

holds for any $\widehat{\boldsymbol{\Theta}}(t) \in \mathcal{H}$. We can replace the term $\widehat{\lambda}_i(t)$ appearing in the denominators of terms in (3.95) with $E\left[L_i^\pi\right]$, and add a term $l_{3i}\left(\widehat{\mathbf{\Theta}}(t)\right)\dot{\widehat{\lambda}}_i(t)$ to account for the change in the expression due to the replacement such that all the terms in $l_{3i}\left(\widehat{\mathbf{\Theta}}\right)$ are bounded. Using the boundedness of terms involved in $l_{3i}\left(\widehat{\mathbf{\Theta}}\right)$, and then using arguments similar to that used in obtaining (3.94), we can show that there is some χ_5 such that

$$\sum_{i\in\mathcal{N}} l_{3i}\left(\widehat{\mathbf{\Theta}}(t)\right)\dot{\widehat{\lambda}}_i(t) \leq \chi_5 d_{8N}\left(\widehat{\mathbf{\Theta}}(t), \widetilde{\mathcal{H}}\right) \tag{3.99}$$

holds for any $\widehat{\mathbf{\Theta}}(t) \in \mathcal{H}$. Thus, we can use the observations in (3.94), (3.98) and (3.99) along with (3.95) to conclude that

$$\begin{aligned}
&\frac{dL\left(\widehat{\mathbf{\Theta}}(t)\right)}{dt} \leq \Delta_1\left(\widehat{\mathbf{\Theta}}(t)\right) + \Delta_2\left(\widehat{\mathbf{\Theta}}(t)\right) \qquad (3.100)\\
&-\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(U_i^E\right)'\left(\widetilde{\widehat{\mu}_i(t)} - U_i^V\left(\widetilde{\widehat{v}_i(t)}\right)\right)\\
&\left(\frac{1}{u_i(t)}\left(m_i^\pi - \widetilde{\widehat{\mu}_i(t)}\right) - \frac{\left(U_i^V\right)'\left(\widetilde{\widehat{v}_i(t)}\right)}{u_i(t)}\left(\frac{E\left[L_i^\pi\left(q_i^\pi\left(F_i^\pi\right) - m_i^\pi\right)^2\right]}{E\left[L_i^\pi\right]} - \widetilde{\widehat{v}_i(t)}\right)\right)\\
&+\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t) d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\frac{1}{u_i(t)}\left(\sigma_i^\pi - \widetilde{\widehat{\sigma}_i(t)}\right)\\
&+\sum_{i\in\mathcal{N}}\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right) - d_i^\pi\right)\frac{1}{u_i(t)}\left(p_i^d\frac{\sigma_i^\pi}{\overline{p}_i} - 1\right)\\
&+\sum_{i\in\mathcal{N}}\widehat{\lambda}_i(t) b_i^\pi\frac{1}{u_i(t)}\left(\sigma_i^\pi - \widetilde{\widehat{\sigma}_i(t)}\right) - \sum_{i\in\mathcal{N}} b_i^\pi\left(\rho_i^\pi - \widetilde{\widehat{\rho}_i(t)}\right)\\
&+\sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{\sigma_i^\pi}{\left(1+\overline{\beta}_i\right)} - \frac{E\left[L_i^\pi\right]\sigma_i^\pi}{u_i(t)}\right)\\
&+\sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{\sigma_i^\pi E\left[L_i^\pi\right]}{u_i(t)} - \frac{\rho_i^\pi}{\tau_{slot}}\right)\\
&+\left(\chi_3 + \chi_4 + \chi_5 - \chi_2\right) d_{8N}\left(\widehat{\mathbf{\Theta}}(t), \widetilde{\mathcal{H}}\right).
\end{aligned}$$

Let $\chi_2 = \chi_3 + \chi_4 + \chi_5 + 1$, and let

$$\begin{aligned} \Delta\left(\mathbf{\Theta}\right) &= \Delta_1\left(\mathbf{\Theta}\right) + \Delta_2\left(\mathbf{\Theta}\right) + \Delta^{(d)}\left(\mathbf{\Theta}\right) + \Delta^{(b)}\left(\mathbf{\Theta}\right) \\ &\quad + \Delta^{(\pi,q)}\left(\mathbf{\Theta}\right) + \Delta^{(\pi,r)}\left(\mathbf{\Theta}\right) + \Delta_3\left(\mathbf{\Theta}\right), \end{aligned} \tag{3.101}$$

where

$$\begin{aligned} \Delta^{(\pi,q)}\left(\widehat{\mathbf{\Theta}}(t)\right) &= -\sum_{i\in\mathcal{N}} \frac{\left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)}{u_i(t)} \left(\left(U_i^E\right)' \left(\widetilde{\widehat{\mu}_i(t)} - U_i^V\left(\widetilde{\widehat{v}_i(t)}\right)\right)\right. \\ &\quad \left(\left(m_i^\pi - \widetilde{\widehat{\mu}_i(t)}\right) - \left(U_i^V\right)'\left(\widetilde{\widehat{v}_i(t)}\right)\left(v_i^\pi - \widetilde{\widehat{v}_i(t)}\right)\right) \\ &\quad \left. + \sum_{i\in\mathcal{N}} d_i^\pi \left(\frac{p_i^d}{\overline{p}_i}\right)\left(\sigma_i^\pi - \widetilde{\widehat{\sigma}_i(t)}\right) + \sum_{i\in\mathcal{N}} b_i^\pi \left(\frac{\sigma_i^\pi}{\left(1+\overline{\beta}_i\right)} - \frac{\widetilde{\widehat{\sigma}_i(t)}}{\left(1+\overline{\beta}_i\right)}\right)\right), \\ \Delta^{(\pi,r)}\left(\widehat{\mathbf{\Theta}}(t)\right) &= -\sum_{i\in\mathcal{N}} b_i^\pi\left(\rho_i^\pi - \widetilde{\widehat{\rho}_i(t)}\right), \\ \Delta^{(d)}\left(\widehat{\mathbf{\Theta}}(t)\right) &= \sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right)\widehat{\lambda}_i(t)\left(h_i^D\left(\widehat{d}_i(t)\right) - d_i^\pi\right)\frac{1}{u_i(t)}\left(p_i^d\frac{\sigma_i^\pi}{\overline{p}_i} - 1\right), \\ \Delta^{(b)}\left(\widehat{\mathbf{\Theta}}(t)\right) &= \sum_{i\in\mathcal{N}} \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{\sigma_i^\pi}{\left(1+\overline{\beta}_i\right)} - \rho_i^\pi\right), \\ \Delta_3\left(\widehat{\mathbf{\Theta}}(t)\right) &= -d_{8N}\left(\widehat{\mathbf{\Theta}}(t), \widetilde{\mathcal{H}}\right), \end{aligned}$$

and recall that

$$\begin{aligned}
\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right) &= -\sum_{i\in\mathcal{N}}\frac{(1+\overline{\beta}_i)}{u_i(t)}E\left[L_i^\pi\left(\phi^Q\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t),F_i^\pi\right),\widehat{\boldsymbol{\theta}}_i(t),F_i^\pi\right)\right.\right.\\
&\quad -\phi^Q\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t),F_i^\pi\right),\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t),F_i^\pi\right)\\
&\quad \left.\left.-\left(U_i^E\right)'\left(\widehat{\mu}_i(t)-U_i^V\left(\widehat{v}_i(t)\right)\right)\left(U_i^V\right)'\left(\widehat{v}_i(t)\right)2\left(\widehat{m}_i(t)-m_i^\pi\right)\left(q_i^*\left(t\right)-\widehat{m}_i(t)\right)\right)\right],\\
\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) &= -\frac{1}{\tau_{slot}}E\left[\phi^R\left(\mathbf{r}^*(t),\widehat{\mathbf{b}}(t),C^\pi\right)-\phi^R\left(\mathbf{r}^\pi\left(C^\pi\right),\widehat{\mathbf{b}}(t),C^\pi\right)\right]\\
&\quad -\sum_{i\in\mathcal{N}}\frac{(1+\overline{\beta}_i)}{u_i(t)}E\left[L_i^\pi\left(\phi^Q\left(q_i^*\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t),F_i^\pi\right),\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t),F_i^\pi\right)\right.\right.\\
&\quad \left.\left.-\phi^Q\left(q_i^\pi\left(F_i^\pi\right),\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t),F_i^\pi\right)\right)\right].
\end{aligned}$$

Hence, we can rewrite (3.100) as follows:

$$\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} \leq \Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right). \tag{3.102}$$

Next, we show that all the functions $\Delta_1(\boldsymbol{\Theta})$, $\Delta_2(\boldsymbol{\Theta})$, $\Delta^{(d)}(\boldsymbol{\Theta})$, $\Delta^{(b)}(\boldsymbol{\Theta})$, $\Delta^{(\pi,q)}(\boldsymbol{\Theta})$, $\Delta^{(\pi,r)}(\boldsymbol{\Theta})$ and $\Delta_3(\boldsymbol{\Theta})$, from the definition (3.101) of $\Delta(\boldsymbol{\Theta})$ are non-positive for $\boldsymbol{\Theta}\in\mathcal{H}^*$ so that $\Delta(\boldsymbol{\Theta})$ is non-positive for all $\boldsymbol{\Theta}\in\mathcal{H}^*$, and that $\Delta(\boldsymbol{\Theta})<0$ for $\boldsymbol{\Theta}\notin\mathcal{H}^*$.

The non-positivity of the functions $\Delta^{(d)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ and $\Delta^{(b)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ follows from complementary slackness conditions for OPTSTAT given in (3.63)-(3.64), and the following observations about the optimal solution to OPTSTAT which follows from the feasibility of the optimal solution (specifically that it satisfies (3.43) and (3.44))

$$p_i^d\frac{\sigma_i^\pi}{\overline{p}_i}\leq 1,\ \ \frac{\sigma_i^\pi}{(1+\overline{\beta}_i)}\leq\rho_i^\pi\ i\in\mathcal{N}.$$

For instance, if we consider the term $\left(h_i^B\left(\widehat{b}_i(t)\right)-b_i^\pi\right)\left(\frac{\sigma_i^\pi}{(1+\overline{\beta}_i)}-\rho_i^\pi\right)$ in the definition of $\Delta^{(b)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$, we see that if $b_i^\pi > 0$, $\frac{\sigma_i^\pi}{(1+\overline{\beta}_i)} = \rho_i^\pi$ due to (3.64), and thus the term is zero. The case for $b_i^\pi = 0$ follows from the feasibility condition (3.44) and the fact that $h_i^B\left(\widehat{b}_i(t)\right) \geq 0$.

Next, we show that $\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ is non-positive and that it is negative unless $\widetilde{\widehat{\boldsymbol{\mu}}(t)} = \mathbf{m}^\pi$ and $\widetilde{\widehat{\mathbf{v}}(t)} = \mathbf{v}^\pi$. Since $\widetilde{\widehat{\boldsymbol{\Theta}}(t)}$ is an element of $\widetilde{\mathcal{H}}$, for each $i \in \mathcal{N}$, there is some feasible quality adaptation policy $\left((q_i(f_i))_{f_i \in \mathcal{F}_i}\right)$ satisfying (see definition of $\widetilde{\mathcal{H}}$ in (3.87))

$$\widetilde{\widehat{\mu}_i(t)} = \frac{E\left[L_i^\pi q_i\left(F_i^\pi\right)\right]}{E\left[L_i^\pi\right]},\ \widetilde{\widehat{v}_i(t)} \geq \operatorname{Var}\left(q_i\left(F_i^\pi\right)\right),\ \widetilde{\widehat{\sigma}_i(t)} \geq \frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right]}{\mathbb{E}\left[L_i^\pi\right]}, \tag{3.103}$$

Using Lemma 3.5 (a), noting that $\left((q_i^\pi(f_i))_{f_i \in \mathcal{F}_i}\right)$ is the unique optimal solution and $\left((q_i(f_i))_{f_i \in \mathcal{F}_i}\right)$ is a feasible solution to the optimization problem considered in Lemma 3.5 (a), we have

$$\begin{aligned}
&U_i^E\left(m_i^\pi - U_i^V\left(v_i^\pi\right)\right) - \sum_{i\in\mathcal{N}} d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\sigma_i^\pi - \sum_{i\in\mathcal{N}}\frac{b_i^\pi}{(1+\overline{\beta}_i)}\sigma_i^\pi \\
&\geq\ U_i^E\left(\frac{\mathbb{E}\left[L_i^\pi q_i\left(F_i^\pi\right)\right]}{\mathbb{E}\left[L_i^\pi\right]} - U_i^V\left(\operatorname{Var}\left(q_i\left(F_i^\pi\right)\right)\right)\right) \qquad (3.104)\\
&\quad -\sum_{i\in\mathcal{N}} d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\left(\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right]}{\mathbb{E}\left[L_i^\pi\right]}\right) - \sum_{i\in\mathcal{N}}\frac{b_i^\pi}{(1+\overline{\beta}_i)}\left(\frac{\mathbb{E}\left[L_i^\pi F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right]}{\mathbb{E}\left[L_i^\pi\right]}\right)\\
&\geq\ U_i^E\left(\widetilde{\widehat{\mu}_i(t)} - U_i^V\left(\widetilde{\widehat{\mathbf{v}}(t)}\right)\right) - \sum_{i\in\mathcal{N}} d_i^\pi\left(\frac{p_i^d}{\overline{p}_i}\right)\widetilde{\widehat{\sigma}_i(t)} - \sum_{i\in\mathcal{N}}\frac{b_i^\pi}{(1+\overline{\beta}_i)}\widetilde{\widehat{\sigma}_i(t)}, \qquad (3.105)
\end{aligned}$$

where the second inequality follows from (3.103). Since $\left((q_i^\pi(f_i))_{f_i \in \mathcal{F}_i}\right)$ is the unique optimal solution, the inequality in (3.104) is strict unless $\left((q_i^\pi(f_i))_{f_i \in \mathcal{F}_i}\right) = \left((q_i(f_i))_{f_i \in \mathcal{F}_i}\right)$. Also, the inequality in (3.105) is strict unless $\widetilde{\widehat{\mu}_i(t)} = \frac{E\left[L_i^\pi q_i\left(F_i^\pi\right)\right]}{E\left[L_i^\pi\right]}$ and $\widetilde{\widehat{v}_i(t)} = \operatorname{Var}\left(q_i\left(F_i^\pi\right)\right)$ for each $i \in \mathcal{N}$. Now, since $U_i^E(.)$ and $-U_i^V(.)$ are concave functions of their arguments, and since $\left(U_i^E\right)'(.)$ is non-negative, we have

$$U_i^E\left(m_i^\pi - U_i^V\left(v_i^\pi\right)\right) \leq U_i^E\left(\widetilde{\widehat{m}_i(t)} - U_i^V\left(\widetilde{\widehat{v}_i(t)}\right)\right)$$
$$+\left(U_i^E\right)'\left(\widetilde{\widehat{\mu}_i(t)} - U_i^V\left(\widetilde{\widehat{v}_i(t)}\right)\right)\left(\widetilde{\widehat{\mu}_i(t)} - m_i^\pi - \left(U_i^V\right)'\left(\widetilde{\widehat{v}_i(t)}\right)\left(v_i^\pi - \widetilde{\widehat{v}_i(t)}\right)\right).$$

By combining the above inequality and (3.105), we have that $\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ is non-positive. Further, since the inequality in (3.104) is strict unless $\left(\left(q_i^\pi\left(f_i\right)\right)_{f_i\in\mathcal{F}_i}\right) = \left(\left(q_i\left(f_i\right)\right)_{f_i\in\mathcal{F}_i}\right)$, and the inequality in (3.105) is strict unless $\widetilde{\widehat{\mu}_i(t)} = \frac{E\left[L_i^\pi q_i\left(F_i^\pi\right)\right]}{E\left[L_i^\pi\right]}$ and $\widetilde{\widehat{v}_i(t)} = \mathrm{Var}\left(q_i\left(F_i^\pi\right)\right)$ for each $i \in \mathcal{N}$, we can conclude $\widetilde{\widehat{\boldsymbol{\mu}}(t)} = \mathbf{m}^\pi$ and $\widetilde{\widehat{\mathbf{v}}(t)} = \mathbf{v}^\pi$.

$$\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0 \text{ only if } \widetilde{\widehat{\boldsymbol{\mu}}(t)} = \mathbf{m}^\pi \text{ and } \widetilde{\widehat{\mathbf{v}}(t)} = \mathbf{v}^\pi. \tag{3.106}$$

Using similar arguments along with Lemma 3.5 (b), we can show that

$$\Delta^{(\pi,r)}\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0 \text{ only if } \widetilde{\widehat{\boldsymbol{\rho}}_i(t)} = \boldsymbol{\rho}^\pi. \tag{3.107}$$

Next we consider the term $\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right)$. Using Lemma 3.3 (c), we can show that $\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right) \leq 0$ and that

$$\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0 \text{ only if } \widehat{\mathbf{m}}(t) = \mathbf{m}^\pi. \tag{3.108}$$

Next we consider the term $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right)$. Using the fact that $q_i^*\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t)}, f_i\right)$ and $\mathbf{r}^*(t)$ are optimal solutions to optimization problems QNOVA$\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t)}, f_i\right)$ and RNOVA$\left(\widehat{\mathbf{b}}(t), c\right)$ respectively, we can conclude that $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) \leq 0$.

Also, note that $\Delta_3\left(\boldsymbol{\Theta}\right) = -d_{8N}\left(\widehat{\boldsymbol{\Theta}}, \widetilde{\mathcal{H}}\right)$ is non-positive, and

$$\Delta_3\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0 \text{ only if } \widetilde{\widehat{\boldsymbol{\mu}}(t)} = \widehat{\boldsymbol{\mu}}(t),\ \widetilde{\widehat{\mathbf{v}}(t)} = \widehat{\mathbf{v}}(t) \text{ and } \widetilde{\widehat{\boldsymbol{\rho}}(t)} = \widehat{\boldsymbol{\rho}}(t). \tag{3.109}$$

Next, we argue that $\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0$ only if

$$\left(\widehat{\boldsymbol{\rho}}(t), \left(h_i^B\left(\widehat{b}_i(t)\right)\right)_{i\in\mathcal{N}}, \left(h_i^D\left(\widehat{d}_i(t)\right)\right)_{i\in\mathcal{N}}\right) \in \mathcal{X}^\pi.$$

Suppose that $\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0$. Then, $\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right) + \Delta^{(\pi,r)}\left(\widehat{\boldsymbol{\Theta}}(t)\right) + \Delta_3\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0$, and from (3.106), (3.107) and (3.109), we can conclude that $\widehat{\boldsymbol{\mu}}(t) = \mathbf{m}^\pi$, $\widehat{\mathbf{v}}(t) = \mathbf{v}^\pi$ and $\widehat{\boldsymbol{\rho}}_i(t) = \boldsymbol{\rho}^\pi$. We also have that $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0$, and hence

$$\begin{aligned} \phi^R\left(\mathbf{r}^*\left(\widehat{\mathbf{b}}(t), c\right), \widehat{\mathbf{b}}(t), c\right) &= \phi^R\left(\mathbf{r}^\pi(c), \widehat{\mathbf{b}}(t), c\right),\ \forall\ c \in \mathcal{C}, \\ \phi^Q\left(q_i^*\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), f_i\right), \widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), f_i\right) &= \phi^Q\left(q_i^\pi(f_i), \widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), f_i\right), \\ &\qquad \forall\ f_i \in \mathcal{F}_i,\ \forall\ i \in \mathcal{N}, \end{aligned}$$

where recall that $\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t) = \left(m_i^\pi, \widetilde{\widehat{\mu}_i(t)}, \widetilde{\widehat{v}_i(t)}, \widehat{b}_i(t), \widehat{d}_i(t)\right)$. Since (from earlier observations in this paragraph) $\widehat{\boldsymbol{\mu}}(t) = \mathbf{m}^\pi$ and $\widehat{\mathbf{v}}(t) = \mathbf{v}^\pi$, we have $\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t) = \left(m_i^\pi, m_i^\pi, v_i^\pi, \widehat{b}_i(t), \widehat{d}_i(t)\right)$. Hence, $\mathbf{r}^\pi(c)$ is an optimal solution to RNOVA$\left(\widehat{\mathbf{b}}(t), c\right)$ for each $c \in \mathcal{C}$, and $q_i^\pi(f_i)$ is an optimal solution to QNOVA$\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), f_i\right)$ for each $f_i \in \mathcal{F}_i$ and $i \in \mathcal{N}$. Hence for each $c \in \mathcal{C}$, $\mathbf{r}^\pi(c)$ satisfies the optimality conditions (3.29)-(3.31) for RNOVA$\left(\widehat{\mathbf{b}}(t), c\right)$. Denote the associated optimal Lagrange multipliers in these conditions by $\left(\chi'(c)\right)_{c\in\mathcal{C}}$ and $\left(\boldsymbol{\omega}'(c)\right)_{c\in\mathcal{C}}$. Similarly, $q_i^\pi(f_i)$ satisfies the optimality conditions (3.32)-(3.34) for QNOVA$\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), f_i\right)$. Let $\left(\left(\gamma_i'(f)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$, $\left(\left(\overline{\gamma}_i(f)'\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$, denote the associated optimal Lagrange multipliers.

Thus, $\left(\left((q_i^\pi(f))_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, (\mathbf{r}^\pi(c))_{c\in\mathcal{C}}\right)$ together with the non-negative constants $\left(\chi'(c)\right)_{c\in\mathcal{C}}$, $\left(\boldsymbol{\omega}'(c)\right)_{c\in\mathcal{C}}$, $\left(\left(\gamma_i'(f)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$ and $\left(\left(\overline{\gamma}_i(f)'\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$ satisfy the optimality conditions for OPTSTAT given in (3.57)-(3.64) with b_i^π replaced by $\widehat{b}_i(t)$ and d_i^π replaced by $\widehat{d}_i(t)$ for each video client $i \in \mathcal{N}$. Note that (3.63)-(3.64) are satisfied since $\Delta^{(d)}\left(\widehat{\boldsymbol{\Theta}}(t)\right) + \Delta^{(b)}\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0$ (since $\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0$). Hence, we have shown that

$$\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right) = 0 \text{ only if } \left(\widehat{\boldsymbol{\rho}}(t), \left(h_i^B\left(\widehat{b}_i(t)\right)\right)_{i\in\mathcal{N}}, \left(h_i^D\left(\widehat{d}_i(t)\right)\right)_{i\in\mathcal{N}}\right) \in \mathcal{X}^{\pi} \quad (3.110)$$

Now, the above discussion along with (3.106), (3.107) (3.109), and (3.110) allow us to conclude that for almost all t

$$\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt} \leq \Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right) \text{ where } \Delta\left(\widehat{\boldsymbol{\Theta}}\right) \leq 0 \;\forall\; \boldsymbol{\Theta} \in \mathcal{H},\; \Delta\left(\boldsymbol{\Theta}\right) < 0 \;\forall\; \boldsymbol{\Theta} \notin \mathcal{H}^*. \quad (3.111)$$

This completes proof of part (a) of the theorem.

Now, we use (3.111) to prove the main claim, i.e., part (b) of the theorem, i.e.,

$$\lim_{t\to\infty} d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \mathcal{H}^*\right) = 0.$$

Suppose

$$\text{limsup}_{t\to\infty} d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t), \mathcal{H}^*\right) = d_0$$

for some $d_0 > 0$. Then, for any $\Delta_t > 0$, there exists (infinite) sequence of increasing numbers $(t_m)_{m\in\mathbb{N}}$ such that $t_1 > \Delta_t$ and for each $m \in \mathbb{N}$, $t_{m+1} - t_m > \Delta_t$ and

$$d_{8N}\left(\widehat{\boldsymbol{\Theta}}\left(t_m\right), \mathcal{H}^*\right) \geq 0.5 d_0.$$

Consider new functions $\overline{\Delta}_1\left(\boldsymbol{\Theta}\right)$, $\overline{\Delta}_2\left(\boldsymbol{\Theta}\right)$, $\overline{\Delta}^{(d)}\left(\boldsymbol{\Theta}\right)$ and $\overline{\Delta}^{(\pi,q)}\left(\boldsymbol{\Theta}\right)$ obtained by replacing $\frac{1}{u_i(t)}$ in the definitions of $\Delta_1\left(\boldsymbol{\Theta}\right)$, $\Delta_2\left(\boldsymbol{\Theta}\right)$, $\Delta^{(d)}\left(\boldsymbol{\Theta}\right)$ and $\Delta^{(\pi,q)}\left(\boldsymbol{\Theta}\right)$ respectively with $\frac{1}{u_{\max}}$ (where $u_{\max}$ is defined in (3.93)). Then, using arguments from the discussion of non-positivity of the functions $\Delta_1\left(\boldsymbol{\Theta}\right)$, $\Delta_2\left(\boldsymbol{\Theta}\right)$, $\Delta^{(d)}\left(\boldsymbol{\Theta}\right)$ and $\Delta^{(\pi,q)}\left(\boldsymbol{\Theta}\right)$, we can show that the new functions give upper bounds, i.e., $\overline{\Delta}_1\left(\boldsymbol{\Theta}\right) \geq \Delta_1\left(\boldsymbol{\Theta}\right)$, $\overline{\Delta}_2\left(\boldsymbol{\Theta}\right) \geq \Delta_2\left(\boldsymbol{\Theta}\right)$, $\overline{\Delta}^{(d)}\left(\boldsymbol{\Theta}\right) \geq \Delta^{(d)}\left(\boldsymbol{\Theta}\right)$ and $\overline{\Delta}^{(\pi,q)}\left(\boldsymbol{\Theta}\right) \geq \Delta^{(\pi,q)}\left(\boldsymbol{\Theta}\right)$. Now, let

$$\overline{\Delta}\left(\mathbf{\Theta}\right)=\overline{\Delta}_1\left(\mathbf{\Theta}\right)+\overline{\Delta}_2\left(\mathbf{\Theta}\right)+\overline{\Delta}^{(d)}\left(\mathbf{\Theta}\right)+\Delta^{(b)}\left(\mathbf{\Theta}\right)+\overline{\Delta}^{(\pi,q)}\left(\mathbf{\Theta}\right)+\Delta^{(\pi,r)}\left(\mathbf{\Theta}\right)+\Delta_3\left(\mathbf{\Theta}\right),$$

so that $\overline{\Delta}\left(\mathbf{\Theta}\right)\geq\Delta\left(\mathbf{\Theta}\right)$. Further, by repeating the arguments above, we can also show that $\overline{\Delta}\left(\widehat{\mathbf{\Theta}}\right)\leq 0,\ \forall\ \mathbf{\Theta}\in\mathcal{H}$, and $\overline{\Delta}\left(\mathbf{\Theta}\right)<0,\ \forall\ \mathbf{\Theta}\notin\mathcal{H}^*$ Note that $\overline{\Delta}\left(\mathbf{\Theta}\right)$ is a continuous function of $\mathbf{\Theta}$, as it is the sum of continuous functions $\overline{\Delta}_1\left(\mathbf{\Theta}\right)$, $\overline{\Delta}_2\left(\mathbf{\Theta}\right)$, $\overline{\Delta}^{(d)}\left(\mathbf{\Theta}\right)$, $\Delta^{(b)}\left(\mathbf{\Theta}\right)$, $\overline{\Delta}^{(\pi,q)}\left(\mathbf{\Theta}\right)$, $\Delta^{(\pi,r)}\left(\mathbf{\Theta}\right)$ and $\Delta_3\left(\mathbf{\Theta}\right)$. It is easy to see that the functions $\overline{\Delta}^{(d)}\left(\mathbf{\Theta}\right)$, $\Delta^{(b)}\left(\mathbf{\Theta}\right)$, $\overline{\Delta}^{(\pi,q)}\left(\mathbf{\Theta}\right)$, $\Delta^{(\pi,r)}\left(\mathbf{\Theta}\right)$ and $\Delta_3\left(\mathbf{\Theta}\right)$ are continuous. The continuity of $\overline{\Delta}_1\left(\mathbf{\Theta}\right)$ and $\overline{\Delta}_2\left(\mathbf{\Theta}\right)$ follows from parts (a) and (c) of Lemma 3.2.

Since $\overline{\Delta}\left(\mathbf{\Theta}\right)$ is a continuous function of $\mathbf{\Theta}$ satisfying $\overline{\Delta}\left(\mathbf{\Theta}\right)<0$ for each $\mathbf{\Theta}\notin\mathcal{H}^*$, we can conclude that

$$\Delta_{\max}:=\max_{\{\mathbf{\Theta}\in\mathcal{H}:d_{8N}(\mathbf{\Theta},\mathcal{H}^*)\geq 0.25d_0\}}\overline{\Delta}\left(\mathbf{\Theta}\right)<0.$$

Since $\widehat{\mathbf{\Theta}}(t)$ is Lipschitz continuous in t (from Lemma 3.6), $d_{8N}\left(\widehat{\mathbf{\Theta}}(t),\mathcal{H}^*\right)$ is also Lipschitz continuous in t so that there can be no abrupt changes in distance of $\widehat{\mathbf{\Theta}}(t)$ from $\mathcal{H}^*$. Thus, we can find some t' such that $d_{8N}\left(\mathbf{\Theta}(t),\mathcal{H}^*\right)\geq 0.25d_0$ for each t in neighborhood $\mathcal{T}_m=[t_m-t',t_m+t']$ of t_m for each m. Further, we pick $\Delta_t>2t'$ so that the sets $\left(\mathcal{T}_m\right)_{m\in\mathbb{N}}$ are disjoint. Then,

$$\int_0^\tau\frac{dL\left(\widehat{\mathbf{\Theta}}(t)\right)}{dt}dt\leq\int_0^\tau\Delta\left(\widehat{\mathbf{\Theta}}(t)\right)dt\leq 2t'\Delta_{\max}m(\tau)$$

where $m(\tau)=\max\left\{m\in\mathbb{N}:t_m+t'<\tau\right\}$. Since $\lim_{\tau\to\infty}m(\tau)=\infty$, we have

$$\lim_{\tau\to\infty}\int_0^\tau\frac{dL\left(\widehat{\mathbf{\Theta}}(t)\right)}{dt}dt=-\infty$$

Thus, we have a contradiction since $\int_0^\tau\frac{dL\left(\widehat{\mathbf{\Theta}}(t)\right)}{dt}dt=L\left(\widehat{\mathbf{\Theta}}(\tau)\right)$ and $L\left(\widehat{\mathbf{\Theta}}(\tau)\right)$ is bounded. This boundedness is due to the continuity of $L(\widehat{\mathbf{\Theta}})$ in $\widehat{\mathbf{\Theta}}$ (see the definition in (3.89)), and due to Lemma 3.6 using which we have that for all τ, $\widehat{\mathbf{\Theta}}(\tau) \in \mathcal{H}$ which is a compact set. Hence, $d_0 = 0$ and thus,

$$\lim_{t\to\infty} d_{8N}\left(\widehat{\mathbf{\Theta}}(t), \mathcal{H}^*\right) = 0.$$

□

3.5.3 Convergence of NOVA and proof of Theorem 3.1

In Subsection 3.5.1, we obtained Theorem 3.2 which says that for almost all sample paths

$$\lim_{S\to\infty}\left(\phi_S\left(\left((q_i^*\left(\boldsymbol{\theta}_i^\pi, f_{i,s}\right))_{i\in\mathcal{N}}\right)_{1\le s\le S}\right) - \phi_S^{opt}\right) = 0,$$

for each $\boldsymbol{\theta}_i^\pi \in \mathcal{H}_i^*$ and each $i \in \mathcal{N}$. This suggests that we can prove the main optimality result for NOVA if we establish convergence of NOVA's parameters to the set $\mathcal{H}_i^*$. The main focus of this subsection is Theorem 3.4 which relates NOVA to the auxiliary differential inclusion (3.75)-(3.82), and thus obtains the desired convergence result for NOVA by using the convergence result obtained in Theorem 3.3 for the differential inclusion. Our approach here relies on viewing the update equations ((3.21)-(3.26) and (3.72)-(3.73)) of NOVA as an asynchronous stochastic approximation update equation (see Chapter 12 of [29] for a detailed discussion on asynchronous stochastic approximation) to relate NOVA to the differential inclusion. After obtaining the convergence result for NOVA in Theorem 3.4, we conclude this section with the proof of Theorem 3.1.

In this subsection, we use the superscript ϵ on NOVA parameters $(m_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(\mu_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(v_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(b_{Q,i,s}^\epsilon)_{i\in\mathcal{N}}$, $(b_{R,i,k}^\epsilon)_{i\in\mathcal{N}}$, $(b_{i,k}^\epsilon)_{i\in\mathcal{N}}$, $(d_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(\lambda_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(\sigma_{i,s}^\epsilon)_{i\in\mathcal{N}}$ and $(\rho_{i,k}^\epsilon)_{i\in\mathcal{N}}$ to emphasize their dependence on ϵ (see NOVA updates in (3.20)-

(3.26)). We refer to the update of NOVA parameters $(m_{i,s_i}, \mu_{i,s_i}, v_{i,s_i}, b_{i,k}, d_{i,s_i}, \lambda_{i,s_i})$ in (3.21)-(3.26) carried out after the selection of segment quality for video client i (following a segment download) as a Q_i-update, and we refer to the update (3.20) on $\mathbf{b}_k$ carried out at the beginning of each slot k as an R-update. Let $\delta\tau^{\epsilon}_{Q,i,s}$ denote the time (in seconds) between the sth and $(s+1)$th Q_i-updates. Let $\delta\tau^{\epsilon}_{R,k}$ denote the time between the kth and $(k+1)$th R-updates, i.e., $\delta\tau^{\epsilon}_{R,k} = \tau_{slot}$ for each k. Let

$$\tau^{\epsilon}_{R,k} = \epsilon \sum_{j=0}^{k-1} \delta\tau^{\epsilon}_{R,j}, \;\; \tau^{\epsilon}_{Q,i,s} = \epsilon \sum_{j=0}^{s-1} \delta\tau^{\epsilon}_{Q,i,j}$$

denote ϵ times the cumulative time for the first k R-updates and s Q_i-updates respectively.

Next, we define time interpolated processes $\left(\widehat{\mathbf{m}}^{\epsilon}(t), \widehat{\boldsymbol{\mu}}^{\epsilon}(t), \widehat{\mathbf{v}}^{\epsilon}(t), \widehat{\mathbf{b}}^{\epsilon}(t), \widehat{\mathbf{d}}^{\epsilon}(t), \widehat{\boldsymbol{\lambda}}^{\epsilon}(t), \widehat{\boldsymbol{\sigma}}^{\epsilon}(t), \widehat{\boldsymbol{\rho}}^{\epsilon}(t)\right)$ associated with NOVA's parameters. For each $i \in \mathcal{N}$ and for $t \in \left[\tau^{\epsilon}_{Q,i,s}, \tau^{\epsilon}_{Q,i,s+1}\right)$, let $\widehat{m}^{\epsilon}_i(t) = m^{\epsilon}_{i,s}$, $\widehat{\mu}^{\epsilon}_i(t) = \mu^{\epsilon}_{i,s}$, $\widehat{v}^{\epsilon}_i(t) = v^{\epsilon}_{i,s}$, $\widehat{b}^{\epsilon}_{Q,i}(t) = b^{\epsilon}_{Q,i,s}$, $\widehat{d}^{\epsilon}_i(t) = d^{\epsilon}_{i,s}$, $\widehat{\lambda}^{\epsilon}_i(t) = \lambda^{\epsilon}_{i,s}$ and $\widehat{\sigma}^{\epsilon}_i(t) = \sigma^{\epsilon}_{i,s}$. Also, for $t \in [k\epsilon, (k+1)\epsilon)$, let $\widehat{b}^{\epsilon}_{R,i}(t) = b^{\epsilon}_{R,i,k}$ and $\widehat{\rho}^{\epsilon}_i(t) = \rho^{\epsilon}_{i,k}$. Recall that $b^{\epsilon}_{Q,i,s}$ and $b^{\epsilon}_{R,i,k}$ are auxiliary variables used in the description of NOVA (in Section 3.4).

For each t, let

$$\begin{aligned}
\widehat{\boldsymbol{\Theta}}^{\epsilon}_{Q}(t) &= \left(\widehat{\mathbf{m}}^{\epsilon}(t), \widehat{\boldsymbol{\mu}}^{\epsilon}(t), \widehat{\mathbf{v}}^{\epsilon}(t), \widehat{\mathbf{b}}^{\epsilon}_{Q}(t), \widehat{\mathbf{d}}^{\epsilon}(t), \widehat{\boldsymbol{\lambda}}^{\epsilon}(t), \widehat{\boldsymbol{\sigma}}^{\epsilon}(t), \widehat{\boldsymbol{\rho}}^{\epsilon}(t)\right), \\
\widehat{\boldsymbol{\Theta}}^{\epsilon}_{R}(t) &= \left(\widehat{\mathbf{m}}^{\epsilon}(t), \widehat{\boldsymbol{\mu}}^{\epsilon}(t), \widehat{\mathbf{v}}^{\epsilon}(t), \widehat{\mathbf{b}}^{\epsilon}_{R}(t), \widehat{\mathbf{d}}^{\epsilon}(t), \widehat{\boldsymbol{\lambda}}^{\epsilon}(t), \widehat{\boldsymbol{\sigma}}^{\epsilon}(t), \widehat{\boldsymbol{\rho}}^{\epsilon}(t)\right),
\end{aligned}$$

Note that definitions $\widehat{\boldsymbol{\Theta}}^{\epsilon}_{Q}(.)$ and $\widehat{\boldsymbol{\Theta}}^{\epsilon}_{R}(.)$ are different only for components $3N+1$ to $4N$. The next result states that for small enough ϵ, the time interpolated versions of NOVA parameters $\widehat{\boldsymbol{\Theta}}^{\epsilon}_{Q}(.)$ and $\widehat{\boldsymbol{\Theta}}^{\epsilon}_{R}(.)$ stay close to the set $\mathcal{H}^*$ (defined in (3.88)) most of the time over long time windows. The proof relies on relating $\widehat{\boldsymbol{\Theta}}^{\epsilon}_{Q}(.)$ and $\widehat{\boldsymbol{\Theta}}^{\epsilon}_{R}(.)$ associated with NOVA to the auxiliary differential inclusion (3.75)-(3.82) by viewing the update equations (3.21)-(3.26) of NOVA as an asynchronous stochastic approximation update equation, and using Theorem 3.3 which states that the differential inclusion converge to the set $\mathcal{H}^*$.

Theorem 3.4. *Let $\widehat{\mathbf{\Theta}}^{\epsilon}_{Q}(0) = \widehat{\mathbf{\Theta}}^{\epsilon}(0) \in \mathcal{H}$. Then, the fraction of time in the time interval $[0,T]$ that $\widehat{\mathbf{\Theta}}^{\epsilon}_{Q}(.)$ and $\widehat{\mathbf{\Theta}}^{\epsilon}_{R}(.)$ spend in a small neighborhood of $\mathcal{H}^*$ converges to one in probability as $\epsilon \to 0$ and $T \to \infty$.*

Proof. This result follows from an extension of Theorem 3.4 in Chapter 12 of [29] which relates asynchronous stochastic approximation (3.21)-(3.26) to its associated differential inclusion (3.75)-(3.82). Theorem 3.4 cannot be directly applied mainly because condition (A3.8) (given in Section 12.3.3, page 418 of [29]) concerning the time between the (asynchronous) updates is not be satisfied in our problem setting (discussed later in the proof). Below, we discuss why Theorem 3.4 can not be directly applied, and an appropriate extension to prove our result. To explain this in more detail, we introduce some notation.

In order to simplify our discussion, we consider the special case of NOVA, NOVA-L1 (described below) which exhibits the key ideas involved in the extension. NOVA-L1 corresponds to a setting with a single video client with U_1^V and U_1^E equal to (linear) identity functions, and no cost constraints. We also assume the allocation in slot k is $r_1^*\left(b^{\epsilon}_{1,k}, c_k\right)$ where $r_1^*\left(b, c\right)$ is a continuous (single valued) function of b for each c (instead of the set valued mapping $\mathcal{R}^*\left(\mathbf{b}_k, c_k\right)$ associated with NOVA). In the single video client case, $r_1^*\left(b, c\right)$ actually does not depend on b. However, below we will not use this property, and will only rely on the continuity of $r_1^*\left(b, c\right)$ (with respect to b) so as to facilitate the extension of the proof to more general settings. To further simplify the notation, we assume that $\overline{\beta}_1 = 0$ and all the segments have the same length l_1. Hence, we need only track $m_{1,.}$ and $b_{1,.}$ since resource allocation and quality adaptation only depend on these parameters (see (3.13) and (3.18)). For this special case, the algorithm NOVA-L1 works as follows:

NOVA-L1.0: Initialize: $m^{\epsilon}_{1,0},\ b^{\epsilon}_{1,0}$. Let $s_1 = 0$.

In each slot $k \geq 0$, carry out the following steps:

RNOVA-L1: At the beginning of slot k, let $b^{\epsilon}_{R,1,k} = b^{\epsilon}_{1,k}$, allocate rate $r^*_1\left(b^{\epsilon}_{1,k}, c_k\right)$ to video client 1, and update $b^{\epsilon}_{1,k}$ as follows:

$$b^{\epsilon}_{1,k+1} \quad = \quad b^{\epsilon}_{1,k} + \epsilon\tau_{slot}. \tag{3.112}$$

QNOVA-L1: In slot k, if video client 1 finishes transmission of the s_1 th segment, let $b^{\epsilon}_{Q,1,s_1+1} = b^{\epsilon}_{1,k+1}$, choose quality $q^*_1\left(\left(m^{\epsilon}_{1,s_1}, b^{\epsilon}_{Q,1,s_1+1}\right), f_{1,s_1+1}\right)$ denoted as q^*_{1,s_1+1} for brevity, and update m^{ϵ}_{1,s_1+1}, $b^{\epsilon}_{1,k+1}$ and s_1 as follows:

$$m^{\epsilon}_{1,s_1+1} \quad = \quad m^{\epsilon}_{1,s_1} + \epsilon\left(q^*_{1,s_1+1} - m^{\epsilon}_{1,s_1}\right), \tag{3.113}$$

$$b^{\epsilon}_{1,k+1} \quad = \quad \left[b^{\epsilon}_{1,k+1} - \epsilon l_1\right]_{\underline{b}}, \tag{3.114}$$

$$s_1 \quad = \quad s_1 + 1.$$

Note that $b^{\epsilon}_{Q,1,s}$ is the value of $b^{\epsilon}_{1,\cdot}$ used in choosing quality for sth segment, and $b^{\epsilon}_{R,1,k}$ is the value of b used in choosing the allocation in kth slot. These are bookkeeping variables, and do not affect the evolution of the algorithm over time. In this proof, we refer to the updates (3.113)-(3.114) in QNOVA-L1 as a Q-update (dropping the subscript '1' used in Section 3.4 since there is just one video client), and the update (3.112) as an R-update.

Let $\delta\tau^{\epsilon}_{Q,s}$ denote the time (in seconds) between the sth and $(s+1)$th Q-updates, i.e., time required to download the sth segment. Let $\delta\tau^{\epsilon}_{R,k}$ denote the time between the kth and $(k+1)$th R-updates. Note that $\delta\tau^{\epsilon}_{R,k} = \tau_{slot}$ for each k whereas

$\delta\tau^{\epsilon}_{Q,s}$ can potentially be different for different s. Let

$$\tau^{\epsilon}_{R,k} = \epsilon \sum_{j=0}^{k-1} \delta\tau^{\epsilon}_{R,j}, \;\; \tau^{\epsilon}_{Q,s} = \epsilon \sum_{j=0}^{s-1} \delta\tau^{\epsilon}_{Q,j}$$

denote ϵ times the cumulative time for the first k R-updates and s Q-updates respectively. We let

$$\begin{aligned}
\tau^{\epsilon}_{R}(t) &= \tau^{\epsilon}_{R,k}, \;\; t \in [k\epsilon, (k+1)\epsilon)\,, \\
\tau^{\epsilon}_{Q}(t) &= \tau^{\epsilon}_{Q,s}, \;\; t \in [s\epsilon, (s+1)\epsilon)\,, \\
\widehat{m}^{\epsilon}_{1}(t) &= m^{\epsilon}_{1,s}, \;\; t \in \left[\tau^{\epsilon}_{Q,s}, \tau^{\epsilon}_{Q,s+1}\right) \\
\widehat{b}^{\epsilon}_{Q}(t) &= b^{\epsilon}_{Q,1,s}, \;\; t \in \left[\tau^{\epsilon}_{Q,s}, \tau^{\epsilon}_{Q,s+1}\right) \\
\widehat{b}^{\epsilon}_{R}(t) &= b^{\epsilon}_{R,1,k}, \;\; t \in \left[\tau^{\epsilon}_{R,k}, \tau^{\epsilon}_{R,k+1}\right)
\end{aligned}$$

Let $\widehat{\mathbf{\Theta}}^{\epsilon}_{Q}(t) = \left(\widehat{m}^{\epsilon}_{1}(t), \widehat{b}^{\epsilon}_{Q}(t)\right)$. Let $\mathcal{F}_{Q,s}$ denote a sigma-algebra that measures at least $m^{\epsilon}_{1,i}, \; b^{\epsilon}_{Q,1,i}, \; \tau^{\epsilon}_{Q,1,i}$ and $F_{1,i}$ for each $i \leq s+1$, and $b^{\epsilon}_{R,1,k}$ and C_k for each $k \leq \frac{\tau^{\epsilon}_{Q,1,s+1}}{\epsilon\tau_{slot}}$. A timing diagram (similar to those given in [29]) is given in Fig. 3.4 illustrating the asynchronous nature of the updates of the variables discussed above.

The stochastic approximation algorithm in our setting is different in two aspects from the one studied in Theorem 3.4 of [29]. First, since the time $\delta\tau^{\epsilon}_{Q,s+1}$ required to download segment $s+1$ depends on the exact instant during the slot in which the segment download begins [2], we do not satisfy the necessary condition (A3.8) (and related assumptions (A3.10) and (A3.14) given in Section 12.3.3, page 418 of [29])) given in Theorem 3.4. More precisely, condition (A3.8) requires that the conditional expectation of $\delta\tau^{\epsilon}_{Q,s}$ with respect to the sigma algebra $\mathcal{F}_{Q,s}$ depends on $\tau^{\epsilon}_{Q,s+1}$ only through the value of $\widehat{\mathbf{\Theta}}^{\epsilon}_{Q}(t)$ at $t = \tau^{\epsilon}_{Q,s+1}$ (and conditions (A3.10)

[2]This requirement may be met in real systems if practical constraints force segment downloads to begin at slot boundaries.

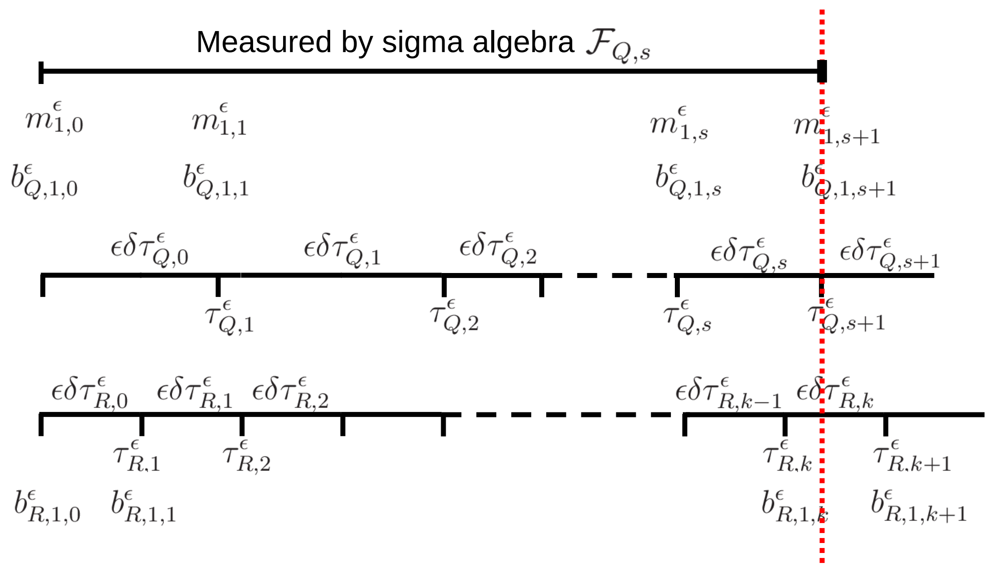

Figure 3.4: NOVA-L1: Asynchronous updates and associated variables

and (A3.14) are related to the continuity properties and averaging behavior of this conditional expectation function). The approach in [29] relies on this assumption to prove the following result (which in turn is used to prove the main result) associated with the component $\tau_Q(t)$ of the weak limits (associated with $\tau^{\epsilon}_Q(t)$):

$$\tau_Q(t) \quad = \quad \int_0^t \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q(\tau_Q(s))\right) ds. \tag{3.115}$$

where for $\mathbf{\Theta} = (m, b)$

$$\overline{u}_Q\left(\mathbf{\Theta}\right) \quad = \quad \tau_{slot} l_1 \frac{E\left[F^{\pi}_1\left(q^*_1\left(\mathbf{\Theta}, F^{\pi}_1\right)\right)\right]}{E\left[r^*_1\left(b, C^{\pi}\right)\right]}.$$

Note that $\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q(\tau_Q(t))\right)$ corresponds to $u_i\left(\widehat{\mathbf{\Theta}}_Q(\tau_Q(t))\right)$ with u_i as defined in (3.83) for the setting in NOVA-L1. Further, we can intuitively see why (3.115) should hold by noting that $\tau_Q(t)$ is roughly the time required to download the first t/ϵ segments (for small ϵ) and viewing $\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q(\tau_Q(s))\right)$ (appearing in the right hand side) as the expected instantaneous segment download time for segment s/ϵ, so that the expression in the right hand side can be viewed as an integral (or roughly the sum) of segment download durations of the first t/ϵ segments.

The first goal of the discussion below is to argue that (3.115) can be established for NOVA-L1. Our argument relies on the fact that when considering time required for the download of a large number of segments, the starting time (and the capacity of the slot associated with that instant) of the download of the first segment makes a negligible impact (unlike the conditional expectation of $\delta\tau^{\epsilon}_{Q,s}$ with respect to $\mathcal{F}_{Q,s}$, considered in condition (A3.8) of [29], which depends on $\tau^{\epsilon}_{Q,s+1}$).

The second aspect which in our setting is different from that in [29] is the fact that the update rules for the $b^{\epsilon}_{Q,1,s}$ and $b^{\epsilon}_{R,1,k}$ are different from those considered in Theorem 3.4 in that they are determined by the evolution of parameter $b^{\epsilon}_{1,k}$ which is *updated on two time scales.* Thus, we must also argue that the weak limits associated with $b^{\epsilon}_{Q,1,s}$ and $b^{\epsilon}_{R,1,k}$ are the same, and that the common limit $\widehat{b}_1(t)$ satisfies (special case of (3.78) for NOVA-L1) given below

$$\dot{\widehat{b}}_1(t) = 1 - \frac{l_1}{u_1\left(\widehat{\mathbf{\Theta}}_Q(t)\right)} + \widehat{z}^b_1\left(\widehat{\mathbf{\Theta}}_Q(t)\right), \tag{3.116}$$

where $\widehat{z}^b_1\left(\widehat{\mathbf{\Theta}}_Q(t)\right) \in -\mathcal{Z}_{\mathcal{H}}\left(\widehat{\mathbf{\Theta}}_Q(t)\right)$. Later in the proof, we show that (3.116) follows once we establish (3.115).

First, let us focus on the proof of (3.115). If we let

$$\begin{aligned} \overline{U}^{\epsilon}_Q(t) &= \epsilon\sum_{i=0}^{\frac{t}{\epsilon}-1} \overline{u}_Q\left(\widehat{\mathbf{\Theta}}^{\epsilon}_Q(\tau^{\epsilon}_{Q,i})\right), \text{ and} \\ \widetilde{U}^{\epsilon}_Q(t) &= \epsilon\sum_{i=0}^{\frac{t}{\epsilon}-1}\left(\delta\tau^{\epsilon}_Q(i) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}^{\epsilon}_Q(\tau^{\epsilon}_{Q,i})\right)\right), \end{aligned} \tag{3.117}$$

then

$$\tau_Q^\epsilon(t) = \epsilon \sum_{i=0}^{\frac{t}{\epsilon}-1} \delta\tau_Q^\epsilon(i) = \overline{U}_Q^\epsilon(t) + \widetilde{U}_Q^\epsilon(t).$$

Next, let

$$W_Q^\epsilon(t) = \tau_Q^\epsilon(t) - \overline{U}_Q^\epsilon(t) = \widetilde{U}_Q^\epsilon(t). \tag{3.118}$$

For ease of reference, we are using notation similar to that used in Chapters 8 and 12 of [29] (and hence we are using redundant notation, for e.g., $W_Q^\epsilon = \widetilde{U}_Q^\epsilon$).

Now, fix t and τ. For any integer p, let $t_i \le t$, $i \le p$, Let $h(.)$ be an arbitrary bounded, continuous and real valued function of its arguments. Hence,

$$\begin{aligned} &E\left[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i \le p\right)\left(W_Q^\epsilon(t+\tau) - W_Q^\epsilon(t)\right)\right] \\ &-E\left[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i \le p\right)\left(\widetilde{U}_Q^\epsilon(t+\tau) - \widetilde{U}_Q^\epsilon(t)\right)\right] = 0. \end{aligned} \tag{3.119}$$

If we show that the expression

$$E\left[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i \le p\right)\left(\widetilde{U}_Q^\epsilon(t+\tau) - \widetilde{U}_Q^\epsilon(t)\right)\right] \tag{3.120}$$

appearing in the above equation goes to zero as $\epsilon \to 0$, then we can use (3.119) along with an approach similar to that in proof of Theorem 2.1 in Chapter 8 of [29] to show that (3.115) holds.

Hence, we next focus on showing that the expression in (3.120) goes to zero as $\epsilon \to 0$. For some fixed $\Delta > 0$, let

$$I_j^{\epsilon,\Delta} = \left\{ i : \frac{j\Delta}{\epsilon} \le i \le \frac{(j+1)\Delta}{\epsilon} \right\}.$$

Then,

$$\begin{aligned}
&\text{limsup}_{\epsilon\to 0} E\left[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i\le p\right)\left(\widetilde{U}_Q^\epsilon(t+\tau) - \widetilde{U}_Q^\epsilon(t)\right)\right] \qquad (3.121)\\
&= \text{limsup}_{\epsilon\to 0} E\Bigg[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i\le p\right)\\
&\qquad \left(\epsilon \sum_{i=\frac{t}{\epsilon}}^{\frac{t+\tau}{\epsilon}-1}\left(\delta\tau_Q^\epsilon(i) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right)\right)\right)\Bigg]\\
&= \lim_{\Delta\to 0}\text{limsup}_{\epsilon\to 0} E\Bigg[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i\le p\right)\\
&\qquad \left(\sum_{j=\frac{t}{\Delta}}^{\frac{t+\tau}{\Delta}-1}\epsilon\sum_{i\in I_j^{\epsilon,\Delta}}\left(\delta\tau_Q^\epsilon(i) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right)\right)\right)\Bigg]\\
&= \lim_{\Delta\to 0}\text{limsup}_{\epsilon\to 0} E\Bigg[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i\le p\right) \qquad (3.122)\\
&\qquad \left(\sum_{j=\frac{t}{\Delta}}^{\frac{t+\tau}{\Delta}-1}\epsilon\mathbb{E}_{\mathcal{F}_{Q,\frac{j\Delta}{\epsilon}}}\left[\sum_{i\in I_j^{\epsilon,\Delta}}\left(\delta\tau_Q^\epsilon(i) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right)\right)\right]\right)\Bigg]
\end{aligned}$$

where the third equality holds since $\mathcal{F}_{Q,\frac{j\Delta}{\epsilon}}$ measures $\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i\le p\right)$ due to the fact that $t_i \le p$ for each $i \le p$ and $j \ge \frac{t}{\Delta}$. Next, we show that (3.122) goes to zero by picking small enough ϵ and δ.

Due to the bounded nature of the quantities involved in the update rules (3.112)-(3.114) for $\widehat{\mathbf{\Theta}}^\epsilon$, we have $\max_{i\in I_j^{\epsilon,\Delta}}\left|\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i+1}^\epsilon) - \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right| = O(\epsilon)$ and hence

$$\max_{i\in I_j^{\epsilon,\Delta}}\left|\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,j}^\epsilon) - \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right| = O(\Delta). \qquad (3.123)$$

Thus, using continuity of $\overline{u}_Q(.)$, we have

$$\lim_{\Delta\to 0} \text{limsup}_{\epsilon\to 0} E\left[\sum_{j=\frac{t}{\Delta}}^{\frac{t+\tau}{\Delta}-1} \epsilon \left|\sum_{i\in I_j^{\epsilon,\Delta}} \left(\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,j}^\epsilon)\right)\right)\right|\right] = 0.$$

Hence, to show that (3.121) is zero (and hence (3.120) goes to zero), it is enough to show that the following term appearing in (3.122) satisfies

$$\lim_{\Delta\to 0} \text{limsup}_{\epsilon\to 0} E\Bigg[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i\leq p\right) \left(\sum_{j=\frac{t}{\Delta}}^{\frac{t+\tau}{\Delta}-1} \epsilon \mathbb{E}_{\mathcal{F}_{Q,\frac{j\Delta}{\epsilon}}}\left[\sum_{i\in I_j^{\epsilon,\Delta}} \left(\delta\tau_Q^\epsilon(i) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,j}^\epsilon)\right)\right)\right]\right)\Bigg] = 0.$$

Let

$$\varsigma_{Q,j}^{\epsilon,\Delta} = \frac{\epsilon}{\Delta} \mathbb{E}_{\mathcal{F}_{Q,\frac{j\Delta}{\epsilon}}}\left[\sum_{i\in I_j^{\epsilon,\Delta}} \delta\tau_Q^\epsilon(i)\right]$$

denote the conditional expectation of the average time (in seconds) for the $\frac{\Delta}{\epsilon}$ updates indexed by the set $I_j^{\epsilon,\Delta}$. Since, for each segment index i and channel slot index k

$$\begin{aligned}
\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon(i\epsilon)\right) &= \left(\widehat{m}_1^\epsilon(\tau_Q^\epsilon(i\epsilon)), \widehat{b}_Q^\epsilon(\tau_Q^\epsilon(i\epsilon))\right) = \left(m_{1,i}^\epsilon, b_{Q,1,i}^\epsilon\right), \\
\widehat{b}_R^\epsilon\left(\tau_R^\epsilon(k\epsilon)\right) &= \widehat{b}_R^\epsilon\left(\tau_{R,k}^\epsilon\right) = b_{R,1,k}^\epsilon.
\end{aligned}$$

we have

$$\sum_{i\in I_j^{\epsilon,\Delta}} l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(i\epsilon\right)\right),F_{Q,i}\right)\right) \geq \sum_{k=\left\lceil\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rceil}^{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)+\left\lceil\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rceil} r^*\left(\widehat{b}_R^\epsilon\left(\tau_R^\epsilon\left(k\epsilon\right)\right),C_k\right), \tag{3.124}$$

and

$$\sum_{i\in I_j^{\epsilon,\Delta}} l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(i\epsilon\right)\right),F_{Q,i}\right)\right) \leq \sum_{k=\left\lfloor\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rfloor}^{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}+2\right)+\left\lfloor\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rfloor} r^*\left(\widehat{b}_R^\epsilon\left(\tau_R^\epsilon\left(k\epsilon\right)\right),C_k\right). \tag{3.125}$$

In the above inequalities, the left hand side is equal to the total size of the segments indexed by the set $I_j^{\epsilon,\Delta}$, and the right hand side is roughly equal to the total allocation over the slots during which these segments are downloaded. The term in the left hand side of (3.124) satisfies

$$\begin{aligned}
&\sum_{i\in I_j^{\epsilon,\Delta}} l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(i\epsilon\right)\right),F_{Q,i}\right)\right) \qquad (3.126)\\
&= \frac{\Delta}{\epsilon}E\left[l_1 F_Q^\pi\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right),F_Q^\pi\right)\right)\right]\\
&+\sum_{i\in I_j^{\epsilon,\Delta}}\left(l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(i\epsilon\right)\right),F_{Q,i}\right)\right)-l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right),F_{Q,i}\right)\right)\right)\\
&+\sum_{i\in I_j^{\epsilon,\Delta}}\left(l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right),F_{Q,i}\right)\right)\right.\\
&\qquad\left.-E\left[l_1 F_Q^\pi\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right),F_Q^\pi\right)\right)\right]\right).
\end{aligned}$$

Similarly, the term in the right hand side of (3.124) satisfies

$$\sum_{k=\left\lceil \frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}} \right\rceil}^{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)+\left\lceil \frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}} \right\rceil} r^*\left(\widehat{b}_R^{\epsilon}\left(\tau_R^{\epsilon}\left(k\epsilon\right)\right), C_k\right) \tag{3.127}$$

$$\geq \left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}} - 2\right) E\left[r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right), C^{\pi}\right)\right]$$

$$-\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}\right) \max_{\left\lceil \frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}} \right\rceil \leq k \leq \left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)+\left\lceil \frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}} \right\rceil}$$

$$\left| r^*\left(\widehat{b}_R^{\epsilon}\left(\tau_R^{\epsilon}\left(k\epsilon\right)\right), C_k\right) - r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right), C_k\right)\right|$$

$$+ \sum_{k=\left\lceil \frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}} \right\rceil}^{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)+\left\lceil \frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}} \right\rceil} \left(r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right), C_k\right) - E\left[r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right), C^{\pi}\right)\right]\right)$$

Using (3.124), (3.126) and (3.127), we have

$$
\begin{aligned}
\Delta\varsigma_{Q,j}^{\epsilon,\Delta} \quad \leq \quad & \Delta\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right)\right)+\Delta E\left[r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),C^{\pi}\right)\right] \\
& \left(2O\left(\frac{\epsilon}{\Delta}\right)+\left|\frac{\sum_{i\in I_j^{\epsilon,\Delta}} l_1F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),F_{Q,i}\right)\right)}{\left(\frac{\Delta}{\epsilon}\right)}\right.\right. \\
& \qquad \left.-E\left[l_1F_Q^{\pi}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),F_Q^{\pi}\right)\right)\right]\right| \\
& +\left(\frac{\varsigma_{Q,j}^{\epsilon,\Delta}}{\tau_{slot}}\right)\left|\frac{\sum_{k=\left\lceil\frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}}\right\rceil}^{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)+\left\lceil\frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}}\right\rceil}\left(r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),C_k\right)\right)}{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)}\right. \\
& \qquad \left.-E\left[r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),C^{\pi}\right)\right]\right| \\
& +\left(\frac{\varsigma_{Q,j}^{\epsilon,\Delta}}{\tau_{slot}}\right)\max_{\left\lceil\frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}}\right\rceil\leq k\leq\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}-2\right)+\left\lceil\frac{\tau_Q^{\epsilon}(j\Delta)}{\epsilon\tau_{slot}}\right\rceil}\left|r^*\left(\widehat{b}_R^{\epsilon}\left(\tau_R^{\epsilon}\left(k\epsilon\right)\right),C_k\right)\right. \\
& \qquad \left.-r^*\left(\widehat{b}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),C_k\right)\right| \\
& +\max_{i\in I_j^{\epsilon,\Delta}}\left|l_1F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(i\epsilon\right)\right),F_{Q,i}\right)\right)\right. \\
& \qquad \left.\left.-l_1F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^{\epsilon}\left(\tau_Q^{\epsilon}\left(j\Delta\right)\right),F_{Q,i}\right)\right)\right|\right).
\end{aligned}
\tag{3.128}
$$

Then, using (3.125) and arguments similar to those above, we have

$$\begin{aligned}
\Delta\varsigma_{Q,j}^{\epsilon,\Delta} \geq\ & \Delta\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right)\right) - \Delta E\left[r^*\left(\widehat{b}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), C^\pi\right)\right] \qquad (3.129)\\
& \left(2O\left(\frac{\epsilon}{\Delta}\right) + \left|\frac{\sum_{i\in I_j^{\epsilon,\Delta}} l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), F_{Q,i}\right)\right)}{\left(\frac{\Delta}{\epsilon}\right)}\right.\right.\\
& \left. - E\left[l_1 F_Q^\pi\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), F_Q^\pi\right)\right)\right]\right|\\
& + \left(\frac{\varsigma_{Q,j}^{\epsilon,\Delta}}{\tau_{slot}}\right)\left|\frac{\sum_{k=\left\lfloor\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rfloor}^{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}+2\right)+\left\lfloor\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rfloor}\left(r^*\left(\widehat{b}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), C_k\right)\right)}{\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}+2\right)}\right.\\
& \left. - E\left[r^*\left(\widehat{b}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), C^\pi\right)\right]\right|\\
& + \left(\frac{\varsigma_{Q,j}^{\epsilon,\Delta}}{\tau_{slot}}\right)\max_{\left\lfloor\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rfloor\leq k\leq\left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}+2\right)+\left\lfloor\frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}}\right\rfloor}\left|r^*\left(\widehat{b}_R^\epsilon\left(\tau_R^\epsilon\left(k\epsilon\right)\right), C_k\right)\right.\\
& \left. - r^*\left(\widehat{b}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), C_k\right)\right|\\
& + \max_{i\in I_j^{\epsilon,\Delta}}\left|l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(i\epsilon\right)\right), F_{Q,i}\right)\right)\right.\\
& \left.\left. - l_1 F_{Q,i}\left(q_Q^*\left(\widehat{\mathbf{\Theta}}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right), F_{Q,i}\right)\right)\right|\right)
\end{aligned}$$

Using the boundedness of the terms, and the fact that $(F_{Q,i})_{i\geq 1}$ and $(C_k)_{k\geq 1}$ are stationary ergodic, the terms appearing in lines 2-5 of (3.128) and (3.129) converge in mean to zero (i.e., we have L^1 convergence to 0). Also, the terms in the last four lines of (3.128) and (3.129) can be made as small as needed by picking a small enough Δ due to the absolute continuity of $q_1^*(.,f)$ and $r_1^*(,c)$ for each f and c. and since we have (3.123) and

$$\max_{\left\lfloor \frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}} \right\rfloor \le k \le \left(\frac{\Delta\varsigma_{Q,j}^{\epsilon,\Delta}}{\epsilon\tau_{slot}}+2\right)+\left\lfloor \frac{\tau_Q^\epsilon(j\Delta)}{\epsilon\tau_{slot}} \right\rfloor} \left|\widehat{b}_R^\epsilon\left(\tau_R^\epsilon\left(k\epsilon\right)\right) - \widehat{b}_Q^\epsilon\left(\tau_Q^\epsilon\left(j\Delta\right)\right)\right| = O(\Delta). \quad (3.130)$$

where the argument for the above property is similar to (3.123) and using the fact that $\varsigma_{Q,j}^{\epsilon,\Delta}$ is bounded by the constant $\delta\tau_{\max}$. Thus, the expression in (3.122) is equal to zero, and consequently (3.120) is also zero. The rest of the proof is similar to that in [29].

In the above arguments, we used the property that the resource allocation $r_1^*\left(b_{1,k}, c_k\right)$ in slot k is a continuous function of $b_{1,k}$ for each c_k. These arguments can be extended if the resource allocation in slot k is picked from $\mathcal{R}^*\left(b_{1,k}^\epsilon, c_k\right)$ where $\mathcal{R}^*\left(b, c\right)$ is an upper semi-continuous set valued map of b taking compact convex values (i.e., $\mathcal{R}^*\left(b, c\right)$ is a convex compact set for each b and $c \in \mathcal{C}$) for each $c \in \mathcal{C}$. For instance, we used (3.123), (3.130) and the continuity of $\overline{u}_Q\left(.\right)$ to argue earlier (see below (3.123)) that

$$\lim_{\Delta\to 0} \text{limsup}_{\epsilon\to 0} E\left[\sum_{j=\frac{t}{\Delta}}^{\frac{t+\tau}{\Delta}-1} \epsilon \left| \sum_{i\in I_j^{\epsilon,\Delta}} \left(\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,j}^\epsilon)\right)\right)\right|\right] = 0.$$

Note that (after relaxing the continuity assumption)

$$\overline{u}_Q\left(\mathbf{\Theta}\right) = \tau_{slot} l_1 \frac{E\left[F_1^\pi\left(q_1^*\left(\mathbf{\Theta}, F_1^\pi\right)\right)\right]}{E\left[r_1^*\left(b, C^\pi\right)\right]}.$$

for $\mathbf{\Theta} = \left(m_1, b_1\right)$ and $r_1^*\left(b_1, c\right) \in \mathcal{R}^*\left(b_1, c\right)$ for each $c \in \mathcal{C}$. Now we can pick

$r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right),c\right) \in \mathcal{R}^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right),c\right)$ for each $c \in \mathcal{C}$ and for each j such that

$$\lim_{\Delta\to 0} \text{limsup}_{\epsilon\to 0} E\left[\sum_{j=\frac{t}{\Delta}}^{\frac{t+\tau}{\Delta}-1} \epsilon \left| \sum_{i\in I_j^{\epsilon,\Delta}} \left(\tau_{slot} l_1 \frac{E\left[F_1^\pi\left(q_1^*\left(\widehat{\boldsymbol{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon), F_1^\pi\right)\right)\right]}{E\left[r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,i}^\epsilon\right), C^\pi\right)\right]} - \tau_{slot} l_1 \frac{E\left[F_1^\pi\left(q_1^*\left(\widehat{\boldsymbol{\Theta}}_Q^\epsilon(\tau_{Q,j}^\epsilon), F_1^\pi\right)\right)\right]}{E\left[r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right), C^\pi\right)\right]} \right) \right| \right] = 0. \tag{3.131}$$

This follows from the following two properties:

(i) the continuity of QR tradeoffs and $q_i^*(.)$, and

(ii) the fact that for small enough Δ, we can always pick $r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right),c\right) \in \mathcal{R}^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right),c\right)$ for each $c \in \mathcal{C}$ and for each j such that

$$\left| \frac{\epsilon}{\Delta} \sum_{i\in I_j^{\epsilon,\Delta}} \frac{1}{E\left[r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,i}^\epsilon\right), C^\pi\right)\right]} - \frac{1}{E\left[r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right), C^\pi\right)\right]} \right|$$

is small enough.

Property (ii) follows from (3.130) and the fact that

$$\overline{\mathcal{R}}_{inv}^*(b) = \left\{ \frac{1}{E\left[r_1^*(b, C^\pi)\right]} : r_1^*(b,c) \in \mathcal{R}^*(b,c) \ \forall\ c \in \mathcal{C} \right\},$$

is an upper semicontinuous set valued map of b taking compact convex values. These properties of $\overline{\mathcal{R}}_{inv}^*(b)$ are essentially consequences of the fact that $\mathcal{R}^*(b,c)$ is an upper semicontinuous set valued map of b taking compact convex values for each c (from Lemma 3.2 (b)). The compactness of $\overline{\mathcal{R}}_{inv}^*(b)$ follows from the compactness of $\mathcal{R}^*(b,c)$ for each c, and the fact that $\overline{\mathcal{R}}_{inv}^*(b)$ is obtained using a continuous map on the elements of $\left(\mathcal{R}^*(b,c)\right)_{c\in\mathcal{C}}$. To show convexity, let $\alpha \in [0,1]$ and consider $x, y \in \overline{\mathcal{R}}_{inv}^*(b)$, i.e., there exists $r_x^*(b,c)$, $r_y^*(b,c) \in \mathcal{R}^*(b,c)$ for each $c \in \mathcal{C}$ such that $x = \frac{1}{E\left[r_x^*(b,C^\pi)\right]}$, $y = \frac{1}{E\left[r_y^*(b,C^\pi)\right]}$. Then, $(\alpha x + (1-\alpha) y) \in \overline{\mathcal{R}}_{inv}^*(b)$

since $(\alpha x+(1-\alpha)y)=\frac{1}{E\left[r^*_{xy}(b,C^\pi)\right]}$ where $r^*_{xy}(b,c)=\alpha' r^*_x(b,c)+(1-\alpha')\,r^*_y(b,c)\in\mathcal{R}^*(b,c)$ (due to convexity of $\mathcal{R}^*(b,c)$), and $\alpha'=\frac{\alpha x}{\alpha x+(1-\alpha)y}\in[0,1]$. Next, we show that $\overline{\mathcal{R}}^*_{inv}(.)$ is an upper semicontinuous set valued map. Note that since $\overline{\mathcal{R}}^*_{inv}(.)$ is uniformly compact, $\overline{\mathcal{R}}^*_{inv}(.)$ is upper semicontinuous if it is closed (see [16] or [44] for a discussion about upper semicontinuous, uniformly compact and closed set valued maps). Consider any sequence $(b_n)_{n\geq 1}$ converging to b, and consider any sequence $(x_n)_{n\geq 1}$ such that $x_n\in\overline{\mathcal{R}}^*_{inv}(b_n)$ for each n, and x_n converges to some x. Then, $\overline{\mathcal{R}}^*_{inv}(b)$ is closed at b if $x\in\overline{\mathcal{R}}^*_{inv}(b)$. Since, $x_n\in\overline{\mathcal{R}}^*_{inv}(b_n)$, there exists $r_{(n)}(c)\in\mathcal{R}^*(b_n,c)$ for each $c\in\mathcal{C}$ such that $x_n=\frac{1}{E\left[r_{(n)}(C^\pi)\right]}$. For any $c\in\mathcal{C}$, we can obtain a convergent subsequence $\left(r_{(n_{k_c})}(c)\right)_{k_c\geq 1}$ that converges to some $r(c)\in\mathcal{R}^*(b_n,c)$ (due to upper semicontinuity of $\mathcal{R}^*(b,c)$). Since $\mathcal{C}$ is finite, we can keep picking subsequences (of subsequences) to obtain a sequence of indices $(n_k)_{k\geq 1}$ such that for each $c\in\mathcal{C}$, $\left(r_{(n_k)}(c)\right)_{k\geq 1}$ converges to some $r(c)\in\mathcal{R}^*(b,c)$. Using this convergence property for each $c\in\mathcal{C}$, and noting that x_n converges to x, we can conclude that $x=\frac{1}{E[r(C^\pi)]}\in\overline{\mathcal{R}}^*_{inv}(b)$. Thus, $\overline{\mathcal{R}}^*_{inv}(.)$ is a closed map, and hence is upper semicontinuous.

Now that we have shown that $\overline{\mathcal{R}}^*_{inv}(b)$ is an upper semicontinuous set valued map of b taking compact convex values, we use this observation to show that property (ii) holds. Due to upper semicontinuity of $\overline{\mathcal{R}}^*_{inv}(b)$ and (3.130), for each $\delta>0$ and $i\in I^{\epsilon,\Delta}_j$, we can find $\overline{r}_{inv,i}\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,i}\right)\right)\in\overline{\mathcal{R}}^*_{inv}\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,i}\right)\right)$ such that $\left|\frac{1}{E\left[r^*_1\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,i}\right),C^\pi\right)\right]}-\overline{r}_{inv,i}\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,i}\right)\right)\right|<\delta$ by picking Δ small enough. Due to convexity of $\overline{\mathcal{R}}^*_{inv}(b)$, there are $r^*_1\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,j}\right),c\right)\in\mathcal{R}^*\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,j}\right),c\right)$ for each $c\in\mathcal{C}$ and for each j that satisfies

$$\frac{1}{E\left[r^*_1\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,j}\right),C^\pi\right)\right]}=\frac{\epsilon}{\Delta}\sum_{i\in I^{\epsilon,\Delta}_j}\overline{r}_{inv,i}\left(\widehat{b}^\epsilon_Q\left(\tau^\epsilon_{Q,i}\right)\right).$$

Then, we see that property (ii) holds since

$$\left| \frac{\epsilon}{\Delta} \sum_{i \in I_j^{\epsilon,\Delta}} \frac{1}{E\left[r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,i}^\epsilon\right), C^\pi\right)\right]} - \frac{1}{E\left[r_1^*\left(\widehat{b}_Q^\epsilon\left(\tau_{Q,j}^\epsilon\right), C^\pi\right)\right]} \right| < \delta$$

for Δ small enough. Since properties (i) and (ii) hold, (3.131) follows. We can similarly extend other arguments in our proof (that relied to continuity $r_1^*(.,c)$) to the case when the resource allocation in slot k is picked from $\mathcal{R}^*\left(b_{1,k}^\epsilon, c_k\right)$ by using the fact that $\mathcal{R}^*(b,c)$ is an upper semi-continuous set valued map of b taking compact convex values for each $c \in \mathcal{C}$.

Now, we focus on proving the result in (3.116) for the weak limit component $\widehat{b}_Q^\epsilon(.)$ associated with time interpolated version $\widehat{b}_Q^\epsilon(.)$ of the parameter $b_{Q,1,s}^\epsilon$. Similar to (3.118), we start by rewriting $\widehat{b}_Q^\epsilon(\tau_Q^\epsilon(t))$ as given below:

$$\widehat{b}_Q^\epsilon(\tau_Q^\epsilon(t)) = b_{1,0}^\epsilon + \widetilde{G}_b^\epsilon(t) + \epsilon \sum_{i=0}^{\frac{t}{\epsilon}-1} \left(\overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right) - l_1\right) + \epsilon \sum_{i=0}^{\frac{t}{\epsilon}-1} Z_{1,i}^\epsilon + E_b^\epsilon(t), \quad (3.132)$$

where

$$\begin{aligned} \widetilde{G}_b^\epsilon(t) &= \epsilon \sum_{i=0}^{\frac{t}{\epsilon}-1} \left(\delta\tau_Q^\epsilon(i) - \overline{u}_Q\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right)\right), \\ E_b^\epsilon(t) &= \epsilon\tau_{slot}\left(\frac{\sum_{i=0}^{\frac{t}{\epsilon}-1}\delta\tau_Q^\epsilon(i)}{\tau_{slot}} - \left\lfloor \frac{\sum_{i=0}^{\frac{t}{\epsilon}-1}\delta\tau_Q^\epsilon(i)}{\tau_{slot}} \right\rfloor\right) \end{aligned}$$

and for each i, $Z_{1,i}^\epsilon$ is the term that accounts for the reflection term associated with the update in (3.114) for the sth segment due to the operator $[.]_{\underline{b}}$. Note that $E_b^\epsilon(t)$ accounts for the fact that R-updates (that increment $b_{1,.}^\epsilon$ by $\epsilon\tau_{slot}$) only occur at slot boundaries which are separated by τ_{slot} seconds. Note that $Z_{1,i}^\epsilon \in -\mathcal{Z}_{\mathcal{H}}\left(\widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_{Q,i}^\epsilon)\right)$ and $|E_b^\epsilon(t)| \leq \epsilon\tau_{slot}$ so that the last term in (3.132) is $O(\epsilon)$.

Now, fix t and τ. For any integer p, let $t_i \leq t$, $i \leq p$, Let $h(.)$ be an arbitrary bounded, continuous and real valued function of its arguments. The proof of (3.116) can be completed just like in [29] (see pages 414-415 and pages 251-257) once we prove that

$$\text{limsup}_{\epsilon\to 0} E\left[h\left(\tau_Q^\epsilon(t_i), \widehat{\mathbf{\Theta}}_Q^\epsilon(\tau_Q^\epsilon(t_i)), i \le p\right)\left(\widetilde{G}_b^\epsilon(t+\tau) - \widetilde{G}_b^\epsilon(t)\right)\right] = 0. \quad (3.133)$$

But, note that $\widetilde{G}_b^\epsilon(t) = \overline{U}_Q^\epsilon(t)$ for each t (defined in (3.117)), and hence, (3.133) holds since we have established this property earlier for the same expression in (3.120).

We have now established that the weak limit $\widehat{b}_Q(.)$ satisfies (3.116). The weak limit components $\widehat{b}_R(.)$ (associated with $b_{R,1,k}^\epsilon$) also satisfies (3.116) since

$$\widehat{b}_Q^\epsilon(t) - \widehat{b}_R^\epsilon(t) = O(\epsilon)$$

which follows from the bounded nature of the increase of $b_{R,1,k}^\epsilon$ due to (3.20), and the fact that the number of slots between two segment download completion instants is bounded.

Thus, we have argued that for the special case NOVA-L1, we can extend Theorem 3.4 of [29] as described in the above discussion. These arguments can be extended to the general setting considered in NOVA also. In particular, we can study NOVA given in (3.21)-(3.26) and (3.72)-(3.73) and relate it to the auxiliary differential inclusion (3.75)-(3.82). We can also show that the following claim in Theorem 3.4 of [29] holds: the fraction of time in the time interval $[0, T]$ that $\widehat{\mathbf{\Theta}}^\epsilon(.)$ and $\widehat{\mathbf{\Theta}}_R^\epsilon(.)$ spends in a small neighborhood of set of limit points of the differential inclusion in (3.75)-(3.82) converges to one in probability as $\epsilon \to 0$ and $T \to \infty$. Now, the main claim of this result now follows from this observation and Theorem 3.3 where we have shown that the set of limit points of the differential inclusion in (3.75)-(3.82) is contained in $\mathcal{H}^*$.

We conclude the proof by verifying that the remaining conditions given in

Theorem 3.4 of [29] are satisfied. Due to Lemma 3.1, we can view NOVA given in (3.21)-(3.26) and (3.72)-(3.73) as a constrained stochastic approximation which satisfies condition (3.1) in Chapter 12 of [29], and hence the set H in the discussion of Theorem 3.4 corresponds to $\mathcal{H}$ in our problem setting. Although the initialization in NOVA, specifically $\mathbf{b}_0$ and $\mathbf{d}_0$, does not ensure that we start in $\mathcal{H}$, we enter and stay in $\mathcal{H}$ in a finite number of slots (as shown in Lemma 3.1), and thus we can view NOVA as a constrained stochastic approximation. Note that the random variables $\xi^{\epsilon}_{s,i}$ considered in the discussion of Theorem 3.4 corresponds to $(F_{i,s}, L_{i,s})$ in our setting, and the condition given in (A3.11) concerning these random variables is clearly satisfied as they take values in a finite set. Condition (A3.1) and (A3.12) can be verified using the boundedness of the quantities associated with these conditions. In particular, note that $0 < \delta\tau_{\min} \leq \delta\tau^{\epsilon}_{Q,s} \leq \delta\tau_{\max} < \infty$ where $\delta\tau_{\min} = \frac{f_{\min} l_{min}}{r_{\max}}$ and $\delta\tau_{\max} = \frac{f_{\max} l_{max}}{\min_{i\in\mathcal{N}} r_{i,\min}}$. The condition (A3.13) is satisfied since, for each $i \in \mathcal{N}$, $(F_{i,s}, L_{i,s})_{s\geq 0}$ is a stationary ergodic process. Conditions (A3.6), (A3.7) and (A3.9) can be verified by letting $\beta^{\epsilon}_{n,\alpha}$ and $\Delta^{\epsilon}_{n,\alpha}$ to be identically zero, and $g^{\epsilon}_{n,\alpha} = \overline{g}_{\alpha}$ where $g^{\epsilon}_{n,\alpha} = \overline{g}_{\alpha}$ are defined based on the right hand side of the differential inclusion in (3.75)-(3.82) taking appropriate care of the terms associated with the update rates. For instance, let $\mathbf{\Theta} = (\mathbf{m}, \boldsymbol{\mu}, \mathbf{v}, \mathbf{b}, \mathbf{d}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathcal{H}$ and let $\boldsymbol{\theta}_i = (m_i, \mu_i, v_i, b_i, d_i)$ for each $i \in \mathcal{N}$, and for each $i \in \mathcal{N}$, we define $\overline{g}_{i+N}$ based on (3.76) as given below:

$$\overline{g}_{i+N}(\mathbf{\Theta}) \quad = \quad \frac{E\left[L^{\pi}_i q^*_i\left(\boldsymbol{\theta}_i, F^{\pi}_i\right)\right]}{\lambda_i} - \mu_i,$$

Note that we have not verified conditions (A3.8), (A3.10) and (A3.14) since our discussion at the beginning of the proof allows us to avoid using these assumptions.

□

We have the following corollary of Theorem 3.4 which says that for small enough ϵ and after running NOVA for long enough, video client i's NOVA parameter stays close to $\mathcal{H}_i^*$(defined in (3.70)) most of the time with high probability.

Corollary 3.1. *Let $\widehat{\mathbf{\Theta}}^\epsilon(0) \in \mathcal{H}$ and $S_\epsilon = \frac{S}{\epsilon}$. Then for each $i \in \mathcal{N}$, the following holds: for any $\delta > 0$, the fraction of segment indices for which $(\boldsymbol{\theta}_{i,s})_{1\leq s\leq S_\epsilon}$ is in a δ-neighborhood of $\mathcal{H}_i^*$ converges to one in probability as ϵ goes to zero and S goes to infinity.*

Proof. The corollary follows by using Theorem 3.4 to conclude that the fraction of time in the time interval $[0, T]$ that $\widehat{\boldsymbol{\theta}}_i^\epsilon(.)$ spends in a small neighborhood of $\mathcal{H}_i^*$ converges to one in probability as $\epsilon \to 0$ and $T \to \infty$. Recall that here $\widehat{\boldsymbol{\theta}}_i^\epsilon(t) = \left(\widehat{m}_i^\epsilon(t), \widehat{\mu}_i^\epsilon(t), \widehat{v}_i^\epsilon(t), \widehat{b}_{Q,i}^\epsilon(t), \widehat{d}_i^\epsilon(t)\right)$. Note that here we are also using the fact that for each video client $i \in \mathcal{N}$, the amount of time between updates is bounded below. $\square$

We have now obtained all the intermediate results required to prove Theorem 3.1 which is given below.

Proof of Theorem 3.1. Part (a) of Theorem 3.1 states that NOVA satisfies the constraints on rebuffering and cost asymptotically, i.e., for each $i \in \mathcal{N}$

$$\text{limsup}_{S\to\infty}\beta_{i,S}\left((q_i^*)_{1:S}, (r_i^*)_{1:K_S}\right) \leq \overline{\beta}_i, \tag{3.134}$$

$$\text{limsup}_{S\to\infty}p_{i,S}\left((q_i^*)_{1:S}\right) \leq \overline{p}_i. \tag{3.135}$$

We first prove that NOVA satisfies rebuffering constraints, i.e., (3.134). Note that

$$\beta_{i,S}\left((q_i^*)_{1:S}, (r_i^*)_{1:K_S}\right) = \frac{\frac{\sum_{s=1}^S l_{i,s} f_{i,s}\left(q_{i,s}^*\right)}{\frac{1}{\tau_{slot} K_S}\sum_{k=1}^{K_S} r_{i,k}^*}}{\sum_{s=1}^S l_{i,s}} - 1. \tag{3.136}$$

Let $T_i(S)$ (measured in seconds) denote the time at which the download of the first S segments of video client i completes. Then, based on NOVA update rules (3.20) and (3.24), and removing the projection operator, we get the following lower bound on the value $b_{Q,i,S}$:

$$\begin{aligned} b_{Q,i,S} &\geq b_{Q,i,0} + \epsilon \left(\frac{\tau_{slot} \left\lfloor \frac{T_i(S)}{\tau_{slot}} \right\rfloor}{\left(1+\overline{\beta}_i\right)} - \sum_{s=1}^{S} l_{i,s} \right) \\ &\geq b_{Q,i,0} - \frac{\epsilon \tau_{slot}}{\left(1+\overline{\beta}_i\right)} + \epsilon \left(\frac{T_i(S)}{\left(1+\overline{\beta}_i\right)} - \sum_{s=1}^{S} l_{i,s} \right). \end{aligned}$$

Hence,

$$\frac{T_i(S)}{\sum_{s=1}^{S} l_{i,s}} \leq \left(1+\overline{\beta}_i\right) + \left(1+\overline{\beta}_i\right) \left(\frac{b_{Q,i,S} - b_{Q,i,0} + \frac{\epsilon \tau_{slot}}{\left(1+\overline{\beta}_i\right)}}{\epsilon \sum_{s=1}^{S} l_{i,s}} \right). \tag{3.137}$$

Now, if we let $K_i(S)$ denote the (random variable associated with) the number of slots which video client i takes to download S segments, then we can express the term appearing in the left hand side of above inequality as

$$\frac{T_i(S)}{\sum_{s=1}^{S} l_{i,s+1}} = \frac{\tau_{slot} \frac{\sum_{s=1}^{S} l_{i,s} f_{i,s}\left(q_{i,s}^*\right)}{\frac{1}{K_i(S)} \sum_{k=1}^{K_i(S)} r_{i,k}^*}}{\sum_{s'=1}^{S} l_{i,s'}} + o(S). \tag{3.138}$$

Now note that any limit point of the sequence $\frac{1}{K_S} \sum_{k=1}^{K_S} r_{i,k}^*$ is also a limit point of the sequence $\frac{1}{K_i(S)} \sum_{k=1}^{K_i(S)} r_{i,k}^*$ since we can uniformly bound $K_i(S) - K_i(S-1)$. Thus, using (3.136), (3.138), (3.137) and the fact that $b_{Q,i,S}$ is bounded (see Lemma 3.1), we can conclude that (3.134) also holds.

Next, we prove that NOVA asymptotically satisfies the cost constraints, i.e., (3.135). Note that the cost per unit video duration associated with the first S segments under NOVA for video client i is

$$p_{i,S}\left(\left(q_i^*\right)_{1:S}\right) = p_i^d \frac{\sum_{s_i=1}^{S} l_{i,s_i} f_{i,s_i}\left(q_{i,s_i}^*\right)}{\sum_{s_i=1}^{S} l_{i,s_i}}. \tag{3.139}$$

Now, using the NOVA update rule (3.25) for parameter d_{i,s_i}, we have

$$d_{i,s_i+1} \geq d_{i,s_i} + \epsilon \left(p_i^d \frac{l_{i,s_i+1} f_{i,s_i+1}\left(q^*_{i,s_i+1}\right)}{\overline{p}_i} - \lambda_{i,s_i} \right).$$

Summing both sides of the above inequality from $s_i = 1$ to S, we have

$$\frac{p_i^d}{\overline{p}_i} \sum_{s_i=1}^{S} l_{i,s_i+1} f_{i,s_i+1}\left(q^*_{i,s_i+1}\right) \leq \sum_{s_i=1}^{S} \lambda_{i,s_i} + \frac{d_{i,S+1} - d_{i,1}}{\epsilon}. \tag{3.140}$$

Next, note that by summing both sides of the NOVA update rule (3.26) for the parameter λ_{i,s_i} from $s_i = 1$ to S, and rearranging the terms, we have

$$\sum_{s_i=1}^{S} \lambda_{i,s_i} = \sum_{s_i=1}^{S} l_{i,s_i+1} - \frac{\lambda_{i,S+1} - \lambda_{i,1}}{\epsilon}.$$

Combining this with (3.140) and dividing by $\sum_{s_i=1}^{S} l_{i,s_i+1}$, we have

$$\frac{p_i^d}{\overline{p}_i} \frac{\sum_{s_i=1}^{S} l_{i,s_i+1} f_{i,s_i+1}\left(q^*_{i,s_i+1}\right)}{\sum_{s_i=1}^{S} l_{i,s_i+1}} \leq 1 - \frac{\lambda_{i,S+1} - \lambda_{i,1}}{\epsilon \sum_{s_i=1}^{S} l_{i,s_i+1}} + \frac{d_{i,S+1} - d_{i,1}}{\epsilon \sum_{s_i=1}^{S} l_{i,s_i+1}}.$$

Now, the result in (3.135) follows from (3.139) and the above inequality by noting that the terms $\lambda_{i,S+1}$, $\lambda_{i,1}$, $d_{i,S+1}$ and $d_{i,1}$ are bounded (from Lemma 3.1), and that $l_{\min} S \leq \sum_{s_i=1}^{S} l_{i,s_i+1} \leq l_{\max} S$.

Next, we prove part (b) of Theorem 3.1 regarding the optimality of NOVA. Using Corollary 3.1 (which says that $(\boldsymbol{\theta}_{i,s})_{1 \leq s \leq S_\epsilon}$ essentially converges to $\mathcal{H}_i^*$) and Lemma 3.2 (a) (which says that $q_i^*(\boldsymbol{\theta}_i, f_i)$ is a continuous function of $\boldsymbol{\theta}_i$), we can conclude that for $\boldsymbol{\theta}_i^\pi \in \mathcal{H}_i^*$

$$\lim_{S \to \infty} \lim_{\epsilon \to 0} \left(\phi_{S_\epsilon} \left(\left(\left(q_i^*\left(\boldsymbol{\theta}_{i,s}, f_{i,s}\right)\right)_{i \in \mathcal{N}} \right)_{1 \leq s \leq S_\epsilon} \right) - \phi_{S_\epsilon} \left(\left(\left(q_i^*\left(\boldsymbol{\theta}_i^\pi, f_{i,s}\right)\right)_{i \in \mathcal{N}} \right)_{1 \leq s \leq S_\epsilon} \right) \right)$$

goes to zero in probability. Now, part (b) of Theorem 3.1 follows from the above observation and Theorem 3.2 which states that for each $i \in \mathcal{N}$ and for almost all sample paths

$$\lim_{S\to\infty}\left(\phi_S\left(\left((q_i^*\left(\boldsymbol{\theta}_i^\pi, f_{i,s}\right))_{i\in\mathcal{N}}\right)_{1\leq s\leq S}\right)-\phi_S^{opt}\right)=0.$$

□

3.6 Extensions

3.6.1 More general QoE models

NOVA can be used for a larger class of QoE models, and still retain its optimality characteristics. For instance, we can consider QoE models such as

$$e_i^S\left(q_i\right)=m_i^{U_i^Q,S}\left(q_i\right)-U_i^V\left(\mathrm{Var}^S\left(q_i\right)\right),$$

where $m_i^{U_i^Q,S}\left(q_i\right)$ is a generalized mean defined as

$$m_i^{U_i^Q,S}\left(q_i\right):=\frac{\sum_{s=1}^S l_{i,s}U_i^Q\left(q_{i,s}\right)}{\sum_{s'=1}^S l_{i,s'}},$$

and U_i^Q is a twice differentiable concave increasing function. We only need two simple modifications to the algorithm NOVA in order to allow for the above QoE model:

1. Modify objective function (3.18) of the optimization problem $\mathrm{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$ associated with the quality adaptation of video client i as given below:

$$\begin{aligned}\phi^{U_i^Q,Q}\left(q_i,\boldsymbol{\theta}_i,f_i\right) &= \left(U_i^E\right)'\left(\mu_i-U_i^V\left(v_i\right)\right)\left(U_i^Q\left(q_i\right)-\left(U_i^V\right)'\left(v_i\right)\left(q_i-m_i\right)^2\right)\\ &\quad-\frac{h_i^B\left(b_i\right)}{\left(1+\overline{\beta}_i\right)}f_i\left(q_i\right)-\frac{p_i^d h_i^D\left(d_i\right)}{\overline{p}_i}f_i\left(q_i\right),\end{aligned}$$

where we have only replaced a q_i term appearing in (3.18) with $U_i^Q(q_i)$.

2. Modify the update rule (3.22) for μ_{i,s_i} as follows:

$$\mu_{i,s_i+1} \quad = \quad \mu_{i,s_i} + \epsilon \left(\frac{l_{i,s_i+1}}{\lambda_{i,s_i}} U_i^Q \left(q^*_{i,s_i+1} \right) - \mu_{i,s_i} \right), \qquad (3.141)$$

so that μ_{i,s_i} keeps track of the generalized mean.

We can show that under the new QoE model, NOVA with the above modifications is still asymptotically optimal, i.e., we can obtain a result similar to Theorem 3.1. This generalization allows us to accommodate QoE models involving generalized mean such as those proposed in [15].

3.6.2 More general channel models

For notational simplicity, we assumed that the network allocation constraint in each slot is a real valued function. However, we can consider more general channel models prevalent in many practical networks like wireless networks using OFDM, where the resource allocation to a video client is the sum of the resource allocation over several sub-resources $w \in \mathcal{W}$ (for e.g., orthogonal subcarriers in OFDM) where $\mathcal{W}$ is a finite set of sub-resources in the network. It is easy to extend the preceding discussion to consider such networks. In particular, we can extend the resource allocation algorithm RNOVA($\mathbf{b}, c$) proposed in NOVA to obtain RNOVA-GC$\big(\mathbf{b}, (c_w)_{w\in\mathcal{W}}\big)$ given below:

$$\begin{aligned}
\max_{\mathbf{r}} \quad & \sum_{i\in\mathcal{N}} h_i^B(b_i)\, r_i \\
\text{subject to} \quad & c_w\left((r_{i,w})_{i\in\mathcal{N}} \right) \leq 0, \ \forall\, w \in \mathcal{W}, \\
& r_i = \sum_{w\in\mathcal{W}} r_{i,w} \geq r_{i,\min} \ \forall\, i \in \mathcal{N}, \\
& r_{i,w} \geq 0, \ \forall\, w \in \mathcal{W}, \ \forall\, i \in \mathcal{N}.
\end{aligned}$$

where the optimization variable $r_{i,w}$ represents the resource allocation to video client i over sub-resource w, and r_i represents the cumulative resource allocation to video client i. If the natural generalization of assumptions on network allocation constraints (e.g., stationary ergodic, Assumption-SF etc) discussed earlier hold, then we can show that the above extension of NOVA (which uses RNOVA-GC$\big(\mathbf{b}, (c_w)_{w \in \mathcal{W}}\big)$ for network resource allocation) is also asymptotically optimal. Similar extensions will typically be possible in general settings where the capacity region can be described using a finite number of convex functions.

3.7 Conclusions

We obtained a simple asymptotically optimal online algorithm NOVA to solve the problem of optimizing video delivery over networks. NOVA is designed to fairly maximize video clients' QoE while taking client preferences on rebuffering time and data costs into account. Further, the distributed, asynchronous and simple nature of NOVA makes it well suited for DASH and current networks. In the next chapter, we study the performance of NOVA using simulations and we discuss the performance of NOVA taking several practical considerations into account.

Chapter 4

NOVA in Practical Networks and Performance Evaluation using Simulation

4.1 Introduction

In this chapter, we discuss the performance of NOVA taking several practical considerations into account and evaluate its performance via simulation. In particular, we focus on the following practical considerations:

- NOVA under other, e.g. legacy, resource allocation policies;
- the performance of quality adaptation in NOVA, referred to as QNOVA, when used for a standalone video client;
- the presence of, and sharing with, other traffic;
- discrete network resources;
- video client implementation considerations for NOVA such as

- discrete levels of quality adaptation,
- choice of ϵ, $(h_i^B(.))_{i\in\mathcal{N}}$ and $(h_i^D(.))_{i\in\mathcal{N}}$,
- reduction of startup delay and frequency of rebuffering,
- playback buffer limits,
- and, video playback pauses;

• and, the performance of NOVA in stochastic networks, i.e., networks with dynamically varying number of video clients.

We consider the above aspects in Sections 4.2-4.5. We evaluate the performance of NOVA using simulation in Section 4.6, and conclude this chapter with a discussion of a possible implementation in Section 4.7.

4.2 NOVA under other resource allocation policies, and QNOVA for a standalone video client

Let us consider the problem of optimizing video delivery in scenarios where we cannot modify or optimize the resource allocation policies, e.g., legacy systems or networks where resource allocation is driven by other considerations. Note in such scenarios, we can still control the quality adaptation at the video clients. In this section, we show that the quality adaptation component in NOVA is still optimal for such scenarios.

Further, we show that QNOVA, i.e., quality adaptation component of NOVA, carries out optimal quality adaptation for a standalone video client.

4.2.1 NOVA under other resource allocation policies

Consider a network where the resource allocation policy cannot be modified or optimized for video delivery. This would be the case in legacy networks where the video clients have to operate under a predetermined resource allocation policy like proportional fair allocation policy (see [30]), or other proprietary (unknown) allocation policies. The following result says that QNOVA is asymptotically optimal for any feasible stationary resource allocation policy $(\mathbf{r}(c))_{c\in\mathcal{C}}$ (see Definition 3.2 of a feasible stationary resource allocation policy in Chapter 3)

Corollary 4.1. *QNOVA is asymptotically optimal for any feasible stationary resource allocation policy* $(\mathbf{r}(c))_{c\in\mathcal{C}}$.

Proof. This result follows once we note that under the given network resource allocation policy $(\mathbf{r}(c))_{c\in\mathcal{C}}$, the offline optimization problem formulation $\mathrm{OPT}_{\mathcal{N}}(S)$, discussed in Section 3.3, breaks into N single video client offline problem formulations $\left(\mathrm{OPT}_{\{i\}}(S)\right)_{i\in\mathcal{N}}$. Recall that $\mathcal{N}$ is the set of video clients considered, and we have added the subscript $\mathcal{N}$ in $\mathrm{OPT}_{\mathcal{N}}(S)$ to emphasize the dependence of the problem formulation on the set of video clients. Thus, for $i \in \mathcal{N}$, the offline optimization problem $\mathrm{OPT}_{\{i\}}(S)$ is obtained by only considering the terms in the objective function and constraints of $\mathrm{OPT}_{\mathcal{N}}(S)$ that involve video client i, and ensuring that the allocation constraints correspond to the fixed resource allocation $(r_i(c))_{c\in\mathcal{C}}$ associated with the video client. □

This result sheds light on an important feature of NOVA that the optimality properties of the adaptation component QNOVA are *insensitive* to the resource allocation in the network as long as the resource allocation policy is stationary. The gains from QNOVA are also explored using simulations in Section 4.6 for a scenario with the legacy proportionally fair resource allocation schemes.

4.2.2 QNOVA for optimizing a standalone video client

We now move away from the network setting, and consider a video client i^* with associated resource allocations $(R_{i^*,k})_{k\geq 1}$ modeled as an exogenous stationary ergodic process (i.e., $R_{i^*,k}$ is the random variable modeling the resource allocation to the client in slot k). This is a reasonable model for a standalone video client accessing video servers in a wide range of scenarios involving wired networks and wireless networks. We have the following important optimality property for QNOVA when used for a standalone video client.

Corollary 4.2. *QNOVA is asymptotically optimal for a standalone video client if the associated resource allocation* $(R_{i^*,k})_{k\geq 1}$ *is an exogenous (i.e., independent of quality adaptation decisions) stationary ergodic process.*

Proof. This result directly follows from the discussion in Chapter 3 by setting $\mathcal{N} = \{i^*\}$ and defining the capacity regions using $(R_{i^*,k})_{k\geq 1}$. □

The above result is significant since the optimization of adaptation in standalone video clients is an important problem in practice, and the result provides optimality guarantees for the solution QNOVA to this problem. Further, this result also reinforces the insensitivity of optimality of the adaptation component in NOVA.

An evolution of the parameters of QNOVA and associated quality adaptation is depicted in Figure 4.1, which also illustrates the response of the quality adaptation in QNOVA to an abrupt capacity drop between time instants 50 secs and 100 secs (see the last subplot). We see that the value of the parameter $b_{i^*,k}$ starts to increase (see the fourth subplot) following the drop, and eventually becomes large enough to force the selection of representations with the least size (see the first subplot).

4.3 NOVA and sharing network resources with other traffic

Video traffic due to stored video is typically carried in networks supporting other types of traffic such as voice traffic, real time video traffic, data traffic due to file

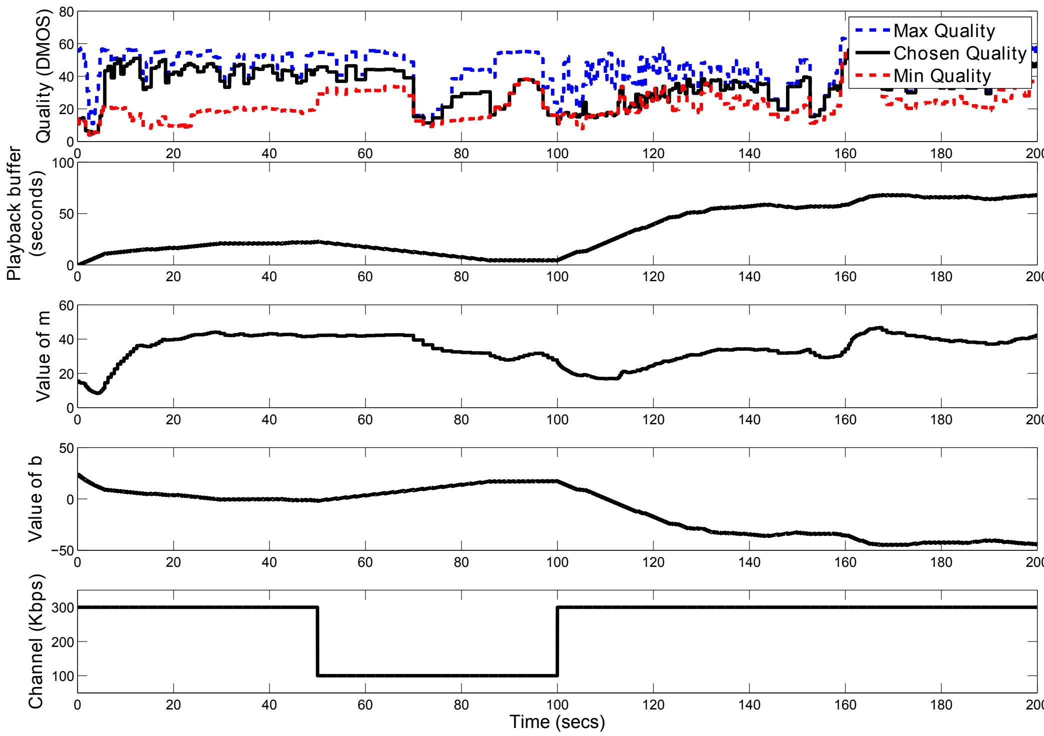

Figure 4.1: Quality adaptation responding to a capacity drop

downloads etc. If the network capacity consumed by other sources of traffic can be modeled as an exogenous stationary ergodic process, then we can extend the optimality result for NOVA given in Theorem 3.1 to this scenario. This can be shown by incorporating the consumption of capacity by other traffic into the stationary ergodic random process $(C_k)_k$ (i.e., the allocation constraints associated with the video clients). This observation about optimality of NOVA is useful as it covers scenarios where the stored video traffic has to compete with other sources of traffic that have higher priority, for e.g., voice traffic and video traffic generated by real time interactive applications like phone calls, video conferencing etc. However, note that in these scenarios, the traffic due to stored video (is low priority to the network and) is only supported using residual capacity in the network.

A more fair approach is to set aside a fixed part of the available network capacity for carrying stored video traffic. This could however lead to inefficient use of network resources if there are too few or too many video clients streaming stored video. A better approach can be obtained by using a more flexible division of network capacity. For instance, we can use the approach developed below which extends NOVA for use in the presence of other data traffic, e.g., associated with users downloading long files. Let $\mathcal{N}_D$ denote the set of data users. The following modification of the resource allocation component in NOVA can be used to this end:

$$\begin{aligned} \max_{((r_i)_{i\in\mathcal{N}\cup\mathcal{N}_D})} \quad & p_V \sum_{i\in\mathcal{N}} h_i^B(b_{i,k}) r_i + \sum_{j\in\mathcal{N}_D} \left(U_j^D\right)' (\rho_{j,k})\, r_j \\ \text{s.t.} \quad & c_k\left((r_i)_{i\in\mathcal{N}\cup\mathcal{N}_D}\right) \leq 0,\ r_i \geq r_{i,\min}\ \forall\, i \in \mathcal{N}\cup\mathcal{N}_D. \end{aligned}$$

The resource allocation in slot k is carried out by obtaining an optimal solution $\left((r^*_{i,k})_{i\in\mathcal{N}\cup\mathcal{N}_D}\right)$ to the above optimization problem, where $(r^*_{i,k})_{i\in\mathcal{N}}$ and $(r^*_{j,k})_{j\in\mathcal{N}_D}$ are respectively the resource allocation to the video clients and data users. Here, U_j^D is a twice differentiable concave function for each data user $j \in \mathcal{N}_D$, and c_k is a convex function characterizing the capacity region in slot k. The parameter $\rho_{j,k}$ tracks the mean resource allocation to data user $j \in \mathcal{N}_D$, and is updated at the beginning of each slot k as follows:

$$\rho_{j,k+1} = \rho_{j,k} + \epsilon\left(r^*_{j,k} - \rho_{j,k}\right).$$

The constant p_V determines the priority given to video clients, i.e., the higher the value of p_V, the higher is the priority given to video over data users.

We can show that NOVA with the above modification in resource allocation is once again asymptotically optimal, i.e., we can establish an optimality result like that given in Theorem 3.1, where we compare the performance of NOVA against that of the optimal offline algorithm solving OPT-D(S), which has objective function

$$p_V \sum_{i\in\mathcal{N}} U_i^E\left(e_i^S\left(q_i\right)\right) + \sum_{j\in\mathcal{N}_D} U_j^D\left(e_j^{K_S,data}\left(r_j\right)\right),$$

and same constraints as in OPT(S) and additional constraints, $r_j \geq r_{j,\min}$ for each $j \in \mathcal{N}_D$, on the resource allocation to the data users. Here, for each data user $j \in \mathcal{N}_D$,

$$e_j^{K_S,data}\left(r_j\right) = m_j^{K_S}\left(r_j\right),$$

represents the QoE for a data user $j \in \mathcal{N}_D$ and is equal to the mean resource allocation to the video client, i.e., this QoE model assumes that the data users care only about their long term time average resource allocation. The objective function of OPT-D(S) also indicates that the constant p_V controls the tradeoff between QoE delivered to video clients versus that delivered to the data users under an optimal solution. Also, note that if there are no video clients (i.e., $\mathcal{N}$ is empty), the above objective corresponds to that of a network carrying out fair resource allocation to the data users where the fairness is implicitly decided by the choice of functions $\left(U_j^D\right)_{j\in\mathcal{N}_D}$. In particular, if U_j^D is log(.) for each $j \in \mathcal{N}_D$, then the resulting allocation is proportionally fair. The above discussion also sheds light on an important property of the resource allocation component in NOVA that it requires only a simple modification of legacy schedulers like proportionally fair schedulers, and hence is well suited for use in current networks.

4.4 NOVA implementation considerations

In this section, we discuss implementation considerations related to NOVA focusing on the resource allocation component of NOVA in Subsection 4.4.1, and the quality adaptation component of NOVA in Subsection 4.4.2.

4.4.1 Discrete network resources

In many practical settings, the set of feasible resource allocations in a slot is *discrete*. For instance, the basic unit of resource allocation in LTE is a Resource Block (RB) which comprises several OFDM sub-carriers for a given time slot, and an RB can be assigned to at most one video client. In such cases, we can use a discrete approximation RNOVA-DISCRETE($\mathbf{b}_k, c_k$) of RNOVA($\mathbf{b}_k, c_k$) given below to obtain the resource allocation in slot k:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \sum_{i \in \mathcal{N}} h_i^B\left(b_{i,k}\right) r_i \\ \text{subject to} \quad & \mathbf{r} \in \mathcal{R}_{c_k}^{discrete}, \end{aligned}$$

where $\mathcal{R}_{c_k}^{discrete}$ is the discrete set of permissible (i.e., permitted by the practical constraints) resource allocation vectors satisfying the allocation constraint c_k for slot k. In many practical settings, allocation constraints are essentially linear, and hence, we can obtain computationally efficient approaches to solve the discrete convex optimization problem RNOVA-DISCRETE($\mathbf{b}_k, c_k$) by exploiting the linearity of the objective and constraint functions. We consider an example of this in Section 4.7 where we obtain the *optimal* solution (and not just a good solution) to the above problem by just finding the maximum of N scalars.

4.4.2 Video client implementation considerations

Discrete quality adaptation

In Chapter 3, we assumed that we have a continuous set of quality choices for each segment. However, in practice, video segments are only available in a *finite* number of representations. Thus, we modify the optimization problem QNOVA$_i(\boldsymbol{\theta}_i, f_i)$

associated with the adaptation in NOVA as follows: the representation chosen for segment s of video client i is the one with quality equal to the optimal solution to QNOVA$_i$-FINITE$(\boldsymbol{\theta}_{i,s}, f_{i,s}, \mathcal{Q}_{i,s})$ given below

$$\begin{aligned} \max_{q_i} \quad & \phi^Q\left(q_i, \boldsymbol{\theta}_{i,s}, f_{i,s}\right) \\ \text{subject to} \quad & 0 \leq q_i \leq q_{\max}, \\ & q_i \in \mathcal{Q}_{i,s}, \end{aligned}$$

where we have modified QNOVA$_i(\boldsymbol{\theta}_{i,s}, f_{i,s})$ by imposing an additional restriction that the quality should be picked from the set $\mathcal{Q}_{i,s}$ of available quality choices for segment s of video client i. For instance, if segment 10 of video client 1 has 4 representations of sizes 200, 300, 500 and 1000 kb with the corresponding quality measurements being DMOS values equal to 38, 48, 62 and 83 respectively, then $\mathcal{Q}_{1,10} = \{38,\ 48,\ 62,\ 83\}$, and $f_{1,10}(38) = 200$ kbps, $f_{1,10}(48) = 300$ kbps etc.

When using QNOVA$_i$-FINITE$(\boldsymbol{\theta}_{i,s}, f_{i,s}, \mathcal{Q}_{i,s})$ for adaptation, extra care is needed while choosing the function U_i^V (which in turn decides the penalty for variability) due to the structure of the objective function

$$\begin{aligned} \phi^Q\left(q_i, \boldsymbol{\theta}_i, f_i\right) \quad = \quad & \left(U_i^E\right)' \left(\mu_i - U_i^V\left(v_i\right)\right) \left(q_i - \left(U_i^V\right)' \left(v_i\right) \left(q_i - m_i\right)^2\right) \\ & - \frac{h_i^B\left(b_i\right)}{\left(1 + \overline{\beta}_i\right)} f_i\left(q_i\right) - \frac{p_i^d h_i^D\left(d_i\right)}{\overline{p}_i} f_i\left(q_i\right), \end{aligned}$$

This is because a very high value of $(U_i^V)'$ could potentially inhibit the above discrete approximation of NOVA's adaptation from selecting representations that correspond to quality choices greater than m_i. This is especially the case when the number of quality choices is small. This happens due to the fact that the term $(q_i - m_i)^2$ could be large for quality choices in the discrete set $\mathcal{Q}_{i,s}$ that are larger than $m_{i,s}$.

Quality adaptation using QNOVA$_i$-FINITE$(\boldsymbol{\theta}_{i,s}, f_{i,s}, \mathcal{Q}_{i,s})$ can be efficiently carried out as it involves a simple task of evaluating the objective function for a few quality choices.

Choice of ϵ, $(h_i^B(.))_{i\in\mathcal{N}}$ and $(h_i^D(.))_{i\in\mathcal{N}}$

The choice of the constant ϵ, and the functions $(h_i^B(.))_{i\in\mathcal{N}}$ and $(h_i^D(.))_{i\in\mathcal{N}}$ used in NOVA can have a significant impact on the convergence and tracking ability of NOVA operating in non-stationary regimes involving short duration videos, discrete quality adaptation (discussed above) etc. Although choosing small ϵ is beneficial for long videos, it can significantly affect the performance (initial transient and tracking ability) of NOVA for short videos. We have observed that choices of ϵ in the range 0.05 to 0.1 typically work well, and often a good choice can be made using trial and error for the system under consideration.

Setting $(h_i^D(.))_{i\in\mathcal{N}}$ as linear functions, with the scaling obtained using trial and error, worked well in our simulations. Good choices of the scaling depend on certain features of QR tradeoffs (of the videos) like their first and second order derivatives (or first and second order differences in the case of discrete quality adaptation). This dependence follows from the dependence of the quality adaptation on the first order derivatives of the QR tradeoffs (see (3.32) of Chapter 3).

Unlike in the case of the choice of $(h_i^D(.))_{i\in\mathcal{N}}$ where (simple) linear functions were enough, we used $h_i^B(.)$ that have the following structure in Section 4.6 (see (4.6)):

$$h_i^B(b_i) = h_{i,0}\left(\frac{b_i}{0.05} + \max\left(\frac{b_i - h_{i,1}}{0.05}, 0\right)^2\right),$$

with carefully chosen constant $h_{i,0}$. Also, note that the constant 0.05 corresponds to ϵ associated with NOVA updates. The linear structure of $h_i^D(d_{i,k})$ was enough to meet the *average* cost constraints in NOVA, whereas the above structure of $h_i^B(b_{i,k})$ allows NOVA to meet average rebuffering constraints (i.e., (3.12) which requires that rebuffering be asymptotically negligible) and a *stronger per-slot requirement* (unlike the 'average requirement') that there is no rebuffering; we explain this below. The constant $h_{i,1}$ is picked to be equal to or (a bit) greater than $b_{i,0} - 20$ (where $b_{i,0}$ is the initialization of the parameter $b_{i,k}$) so that $h_i^B(b_i)$ increases more quickly (i.e., quadratically) when $\frac{b_i}{\epsilon}$ is close to $\frac{b_{i,0}-20}{\epsilon}$, and is very large when $\frac{b_i}{\epsilon}$ is close to $\frac{b_{i,0}-5}{\epsilon}$ so as to force QNOVA to select lower quality representations. This feature of QNOVA is desirable since (we have argued the following in Section 3.4 of Chapter 3 after presenting the algorithm NOVA)

$$\frac{b_{i,k}}{\epsilon} - \frac{b_{i,0}}{\epsilon} \approx \left(\frac{k\tau_{slot}}{(1+\overline{\beta}_i)} - \text{Duration of video downloaded till now} \right).$$

For instance, if $\overline{\beta}_i = 0$ (i.e., video client i prefers not to see any rebuffering), then our choice of $h_{i,1}$ ensures that $h_i^B(b_i)$ is large when the playback buffer has video content of duration less than 5 seconds (i.e., $\frac{b_i}{\epsilon}$ is close to $\frac{b_{i,0}-5}{\epsilon}$) so that QNOVA starts to pick lower (if not lowest) quality representations. This was our main motivation for using $h_i^B(b_{i,k})$ with the above structure, and setting $h_{i,1} \approx b_{i,0} - 20$.

Good choices of the constant $h_{i,0}$ depend again on the characteristics of the QR tradeoffs like the first and second order derivatives (as in the case of good scaling constants for $(h_i^D(.))_{i\in\mathcal{N}}$), and we obtained them via trial and error for the system under consideration.

Reducing startup delay and the frequency of rebufffering

Video client optimization also requires attention to issues like reduction of frequency of rebuffering events and playback startup delay. Frequency of rebuffering events can be reduced by forcing the video client to delay the resumption of playback after a rebuffering event until there is sufficient amounts of video content in the playback buffer.

We can reduce the start up delay by appropriately choosing the initial conditions. For instance, we can pick large values for $b_{i^*,0}$, and small values for $m_{i^*,0}$ (recall that $b_{i^*,0}$ and $m_{i^*,0}$ denote the initialization of parameters $b_{i^*,k}$ and $m_{i^*,s}$ used in NOVA) to encourage selection of representations with smaller size so that they are downloaded quickly at the beginning. An evolution of NOVA's parameters using such an initialization is depicted in Figure 4.1 where we see that, in the beginning, large $b_{i^*,0}$ and small $m_{i^*,0}$ encourages the selection of representations with lower STQ and thus, smaller size.

Playback buffer limits

In practical systems, we might have to operate NOVA under an additional constraint on the size of the playback buffer. Hardware limitations on memory could be a reason for this constraint although the latest smartphones, tablets, laptops etc have plenty of memory and memory limitations are no longer a major concern in the design of video clients. However, it is interesting to note that even data delivery cost considerations can force us to impose this constraint especially when the chances of video client abandonment are high, i.e., the possibility of a video client terminating the video playback without viewing the entire video is high. For instance, suppose that we do not impose any constraint on the size of playback buffer, and hence there are no constraints on the amount of video data downloaded by the video client that has not been viewed yet. Then, if the video client receives very high resource allocation (that is well above even the largest compression rates of available representations), then QNOVA will aggressively download the video segments exploiting the good resource allocation. But, all the downloaded data would be wasted if there is a video client abandonment. In such a scenario, the video client and/or content provider might have to pay for the delivery of these wasted (i.e., not viewed) segments. Thus, high data delivery costs and video client abandonment concerns would motivate the use of playback buffer limits for some types of content.

Note that there is an interesting tradeoff between the size of the playback buffer (small buffers might reduce wastage of video data in the event of a video client abandonment) and the ability of QNOVA (and other adaptation algorithms) to exploit periods of good resource allocation. For instance, consider a user running a video client on a mobile that is moving away from a base station in a cellular network so that the user initially sees better wireless channels but they are declining. Imposing a playback buffer limit will adversely impact the ability of the video client to exploit high initial resource availability. Therefore, we conclude that it may be useful to impose limits on playback buffer size, and this has to be carefully chosen after taking into account the data delivery costs, and the possibility of video client abandonment etc.

In the presence of limits on the size of playback buffer, we can modify QNOVA to slow down the rate of segment download when capacity is abundant. A simple modification would be to stop segment download requests once a playback buffer limit is reached. A better modification would use a 'smoother' approach where we keep reducing the segment download rate as we approach the playback buffer limit, for e.g., we could force QNOVA to delay the next segment download request by a duration proportional to

$$\max\left(\frac{1}{PB_{lim} - PB_{cur}} - \frac{1}{0.5PB_{lim}}, 0\right)$$

where PB_{lim} is the playback buffer limit and PB_{cur} is the current state of the playback buffer. This would ensure that QNOVA slows segment download rate once the playback buffer is large enough, (i.e., greater than $0.5PB_{lim}$), and stops once it reaches the limit.

Note that under NOVA, the issue of wastage of data under video client abandonment is mainly relevant in scenarios where the network resource allocation is exceptionally high. This is due to the fact that QNOVA will switch to higher quality representations under high resource allocation, and hence, excessive buffering of video data can occur only when average resource allocation is consistently higher than average compression rate of the largest representations (or if the video client has set $\overline{\beta}_i$ in NOVA less than zero).

Video playback pauses

If a video client $i \in \mathcal{N}$ pauses the playback of a video, then we stop the use of update rule (3.20) which increments the value of the variable $b_{i,k}$, and resume the use of the update rule when the video client resumes its playback. Recall that (see Section 3.4) $b_{i,k}$ serves an indicator of risk of violation of rebuffering constraints of video client i, and a large value of $b_{i,k}$ would force the selection of low quality representations by the video client and ensure higher priority in resource allocation to the video client. Hence, by temporarily pausing the use of update rule (3.20), we ensure that we are not unnecessarily forcing the video client to lower its quality or forcing the network controller to give higher priority to a paused video client.

We can use the same idea of temporarily pausing the use of update rule (3.20) when the content provider inserts ads during the playback of a video.

4.5 NOVA in stochastic networks

Till now, we focused on networks with a static number of video clients (since the set $\mathcal{N}$ considered in Chapter 3 is a fixed set) and data users. An important feature of real world networks will be dynamics in the number of video clients. Motivated by this, in this section, we study stochastic networks where video clients arrive into the network, utilize network resources to stream video content, and depart.

We start by exploring some of the new challenges associated with video delivery optimization problem in stochastic networks. Recall that we formulated the video delivery optimization problem in Chapter 3 as an optimization problem OPT-BASIC. Similarly, we could formulate video delivery optimization problem in stochastic networks as the following 'stochastic' extension OPT-BASIC-STOCH associated with a time window of K slots:

$$\begin{aligned} \max \quad & \frac{1}{|\mathcal{N}[0,K]|} \sum_{i \in \mathcal{N}[0,K]} U_i^E \left(\text{Mean Quality}_i - \text{Quality Variability}_i\right) \\ \text{subject to} \quad & \text{Rebuffering}_i,\ \text{Cost}_i, \text{ and Stochastic Network constraints,} \end{aligned}$$

where $\mathcal{N}[0,K]$ is the set of video clients who arrive into the network during the K slots. Note that the new objective is the average of U_i^E of the QoE of video clients utilizing the network resources during the K slots, where the functions U_i^E implicitly decide fairness in the delivery of QoE to the video clients. The *stochastic* network constraints reflect the network resource allocation constraints associated with the time varying number of video clients in different slots, and this video client dynamics introduces another potential source of variability in network resource allocation.

Solving OPT-BASIC-STOCH presents new challenges due to the video client dynamics considered in the formulation. Firstly, note that the time duration (i.e., the number of slots) that a video client spends in the network can depend on the resource allocation to it. For instance, a video client receiving high resource allocation over many slots could leave the network early after downloading all the segments of its video before the completion of video playback. However, even if we ignore this dependence on resource allocation (this dependence will be negligible if there are sufficient choices of representations for segments and the quality adaptation algorithm responds to very high or very low resource allocation by appropriately adjusting the effective rate of download of segments), there are issues related to video client dynamics that need to be carefully tackled. For instance, consider a video client i^* that arrives into the network at time slot $\underline{k}$ and leaves the network at time slot $\overline{k}$.

Now, let $\mathcal{N}_1$ denote the set of video clients in the network after the arrival of video client i^* in slot $\underline{k}$, and we refer to this set as the state of the network in slot $\underline{k}$. Suppose that the state of the network does not change until slot $0.5\left(\overline{k}+\underline{k}\right)$, and (the arrival/departure of another video client causes) the network state changes to $\mathcal{N}_2$ in slot $0.5\left(\overline{k}+\underline{k}\right)$ (for simplicity, assume that this is an integer), and suppose that the network state remains the same until slot $\overline{k}$ (when video client i^* leaves). Suppose the video client downloads S_1 segments until slot $0.5\left(\overline{k}+\underline{k}\right)$, and S_2 segments in the remaining slots. Then,

$$\begin{aligned}\mathrm{Var}\left(\left(q_{i^*,s}\right)_{1:(S_1+S_2)}\right) &= \frac{S_1\mathrm{Var}\left(\left(q_{i^*,s}\right)_{1:(S_1)}\right)}{S_1+S_2}+\frac{S_2\mathrm{Var}\left(\left(q_{i^*,s}\right)_{(S_1+1):(S_1+S_2)}\right)}{S_1+S_2}\\ &+\frac{S_1\left(\mathrm{Mean}\left(\left(q_{i^*,s}\right)_{1:(S_1)}\right)-\mathrm{Mean}\left(\left(q_{i^*,s}\right)_{1:(S_1+S_2)}\right)\right)^2}{S_1+S_2}\\ &+\frac{S_2\left(\mathrm{Mean}\left(\left(q_{i^*,s}\right)_{(S_1+1):(S_1+S_2)}\right)-\mathrm{Mean}\left(\left(q_{i^*,s}\right)_{1:(S_1+S_2)}\right)\right)^2}{S_1+S_2},\end{aligned}$$

where

$$\mathrm{Mean}\left(\left(q_{i^*,s}\right)_{1:(S)}\right)=\frac{1}{S}\sum_{s=1}^{S}q_{i^*,s},$$

$$\mathrm{Var}\left(\left(q_{i^*,s}\right)_{1:(S)}\right)=\frac{1}{S}\sum_{s=1}^{S}\left(q_{i^*,s}-\mathrm{Mean}\left(\left(q_{i^*,s}\right)_{1:(S)}\right)\right)^2.$$

From the above expression, we see that $\mathrm{Var}\left(\left(q_{i^*,s}\right)_{1:(S_1+S_2)}\right)$ (and hence the QoE) of the video client i^* is a complex function of resource allocation and quality adaptation associated with the various network states seen during its stay in the network, for e.g., note that S_1 and S_2 depend on the resource allocation to the video client i^* during $\left[\underline{k}, 0.5\left(\overline{k}+\underline{k}\right)\right]$ and $\left[0.5\left(\overline{k}+\underline{k}\right), \overline{k}\right]$ respectively which in turn depend on the network states $\mathcal{N}_1$ and $\mathcal{N}_2$ respectively associated with these time windows. This observation suggests that a direct extension of the approach in Chapter 3, which is based on considering a static network (or a *single* network state), is difficult.

However, this can be done under additional assumptions on the nature of video client dynamics and its impact on the stochastic network constraints, and this is the focus of most of the rest of this section.

Although a theoretical analysis of the performance of NOVA in stochastic networks is difficult without additional simplifying assumptions, the following three features of NOVA suggest that we can expect it to perform well in stochastic networks also.

1. **Adaptation in NOVA is optimal and insensitive to resource allocation:** This property, explored in Section 4.2, essentially guarantees that the adaptation in NOVA is optimal as long the resource allocation can be modeled as a stationary ergodic process. Thus, if the video client dynamics results in a resource allocation to the video clients which is stationary ergodic, the *quality adaptation* in NOVA will perform well. However, this argument cannot be extended to argue optimality of *resource allocation* under video client dynamics.

2. **The tracking ability of NOVA:** Although NOVA was studied for scenarios where a fixed set of video clients see stationary variations in capacity and video QR tradeoffs, NOVA has tracking ability built into it which allows it to perform well in non-stationary regimes. Such non-stationary regimes could include networks with video client dynamics, settings where the video and/or channels exhibit non-stationary behavior etc. The tracking ability follows from the structure of the update rules for the parameters used in NOVA where the current decision is weighted by ϵ. For instance, consider the update rule (3.21) for NOVA parameter m_{i,s_i} repeated below:

$$m_{i,s_i+1} = m_{i,s_i} + \epsilon \left(U_i^E\right)' \left(\mu_i - U_i^V(v_i)\right) \left(U_i^V\right)'(v_i) \left(\frac{l_{i,s_i+1}}{\lambda_{i,s_i}} q^*_{i,s_i+1} - m_{i,s_i}\right).$$

Recall that the parameter m_{i,s_i} is responsible for tracking the mean quality of video client i. This rule keeps updating the value of the estimate of mean based on the quality q^*_{i,s_i+1} of the current segment under consideration. The choice of ϵ decides the impact of the current quality q^*_{i,s_i+1}. Similarly, we can identify the impact of current decisions on other NOVA parameters. This influence of current information allows NOVA to track and adapt in non-stationary settings. Further, we can control the tracking ability by controlling ϵ since we can increase the impact of current decisions on the update, and thus on the tracking ability, by increasing ϵ. However, note that an excessively high value of ϵ can degrade the performance of the algorithm, as NOVA parameters will not be able to converge due to their evolution being swayed by even small changes in the network.

3. **Optimal for static setting:** NOVA comes with strong optimality guarantees for a static setting with a fixed set of video clients. This, along with the tracking ability discussed above, suggests that we can expect good performance under certain assumptions on the video client dynamics. For instance, this would be the case if the video client dynamics were 'slow'. We discuss this in more detail below.

We start with a discussion about an extension of the model in Chapter 3 to a stochastic setting. We index the video clients in the order of their arrival to the network using the variable $j \in \mathbb{N}$. Let $\mathcal{N}(k)$ denote the set of indices of video clients supported by the network during slot k, and we refer to this as the state of the network in slot k. Also, let $N(k) = |\mathcal{N}(k)|$. Let $\mathcal{H}^*_{\mathcal{N}(k)}$ denote the set of optimal NOVA parameters defined in (3.88) associated with network state $\mathcal{N}(k)$. To see some of the issues arising in a dynamic setting, consider Fig. 4.2 which illustrates a typical evolution of distance $d\left(\widehat{\Theta}^{NOVA}(k), \mathcal{H}^*_{\mathcal{N}(k)}\right)$ between current value of NOVA parameters $\widehat{\Theta}^{NOVA}(k)$ to the current set of optimal NOVA parameters $\mathcal{H}^*_{\mathcal{N}(k)}$.The time slots indices $(k_j)_{j\geq 1}$ indicated in the figure correspond to time slots in which there is change in the state of the network due to an arrival or departure of a video client. Recall that Theorem 3.4 of Chapter 3 ensures that $\widehat{\Theta}^{NOVA}(k)$ converges

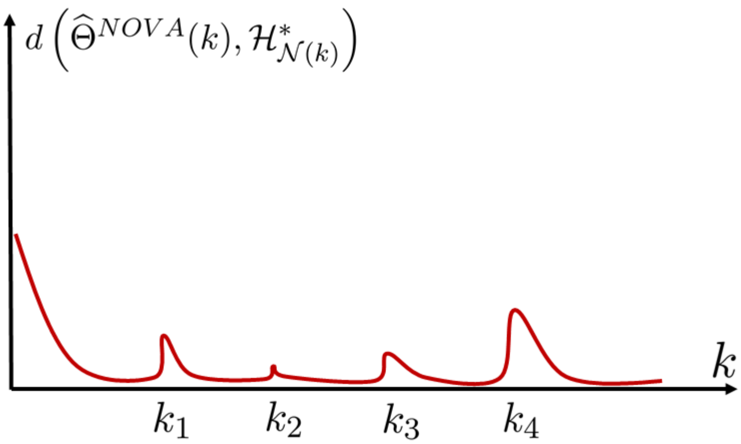

Figure 4.2: A typical evolution of distance of NOVA parameters $\widehat{\Theta}^{NOVA}(k)$ from optimal set $\mathcal{H}^*_{\mathcal{N}(k)}$ of parameters under video client dynamics

to a neighborhood of $\mathcal{H}^*_{\mathcal{N}(k)}$ over time. However, a change in the state of the network will typically result in a new optimal set of parameters, and this causes the abrupt increase in $d\left(\widehat{\Theta}^{NOVA}(k), \mathcal{H}^*_{\mathcal{N}(k)}\right)$ at time slots $(k_i)_{i\geq 1}$. This increase in the 'distance from optimality' will be less pronounced if variations in the network state are typically small, and if the impact of small variations in network state on performance of NOVA is small. We explore this property in the rest of this section, and identify scenarios where we can expect this to hold.

To simplify the notation in the rest of this section, we assume that the segment lengths of all the video clients are fixed and are equal to l. The arguments in this section can be extended to settings with heterogeneous variable-length segments. In the rest of this section, we impose an additional constraint on the structure of the allocation constraint. Specifically, in each slot k, the resource allocation to the current set of video clients $\mathcal{N}(k)$ must satisfy

$$\sum_{j \in \mathcal{N}(k)} c_j(r_j) \leq \overline{c}, \tag{4.1}$$

where $\overline{c}$ is a positive constant representing the cumulative capacity in the system, c_j is a function picked from a finite set $\mathcal{C}_u$ of convex functions where each element is a function that maps video client resource allocation to its impact on the allocation constraint. We refer to these as video client capacity consumption functions. Thus, we are assuming that the allocation constraints are video client-separable, i.e., the impact of capacity consumption of a video client on the allocation constraint is additive. This additivity is satisfied if the network resource (in a slot) comprises orthogonal sub-resources (like in TDMA, FDMA, OFDMA etc) and is typically not satisfied if we do not have orthogonality (for instance, see Gaussian Multiple Access Channels with joint decoding in [13]). Later, we use this separability along with assumptions on independence of evolution of capacity functions to argue that the cumulative capacity consumption of a large number of video clients scales approximately linearly in their number. We also assume that $\mathcal{C}_u$ satisfies (natural modification of) the assumptions on allocation constraint functions discussed in Section 3.2, and in particular, they are convex increasing functions. Also, let $\mathcal{F}$ denote a finite set of QR tradeoffs such that each QR tradeoff picked from this set satisfies the assumptions made on QR tradeoffs in Section 3.2. In particular, each QR tradeoff is a convex increasing function.

To facilitate the analysis of the stochastic setting, we divide the set of video clients associated with the network into classes. Let $\mathcal{M}$ denote the set of all classes of video clients, and we assume that this set is finite. Let $m(j)$ denote the class of the video client with index j. All the video clients of a given class have the same video client preferences and QoE model so that $\overline{\beta}_{j_1} = \overline{\beta}_{j_2}$, $\overline{p}_{j_1} = \overline{p}_{j_2}$ and $U^V_{j_1} = U^V_{j_2}$ for any video client indices j_1 and j_2 such that $m(j_1) = m(j_2)$. Also, $U^E_{j_1} = U^E_{j_2}$

for any video client indices j_1 and j_2 such that $m(j_1) = m(j_2)$ so that all the video clients of a given class are subject to same fairness considerations.

The probability law for the evolution of QR tradeoffs and the video client capacity consumption functions associated with a video client depends only on its class. Associated with each class $m \in \mathcal{M}$ are two probability distributions $\left(\pi_m^{\mathcal{F}}(f)\right)_{f\in\mathcal{F}}$ and $\left(\pi_m^{\mathcal{C}_u}(c)\right)_{c\in\mathcal{C}_u}$ defined on the sets $\mathcal{F}$ and $\mathcal{C}_u$. The evolution $\left(F_{j,s}\right)_{s\geq 0}$ of QR tradeoffs of the jth video client is modeled as a stationary ergodic process with marginal distribution $\left(\pi_{m(j)}^{\mathcal{F}}(f)\right)_{f\in\mathcal{F}}$. Let c_{jk} denote the video client capacity consumption function for slot k associated with video client with index j. The evolution $\left(C_{jk}\right)_k$ of video client capacity consumption functions of the jth video client (of class $m(j)$) is modeled as a stationary ergodic process with marginal distribution $\left(\pi_{m(j)}^{\mathcal{C}_u}(c)\right)_{c\in\mathcal{C}}$.

Let $\mathcal{N}_m(k)$ denote the set of indices associated with the video clients of class $m \in \mathcal{M}$ present in the network during slot k, and let $N_m(k) = |\mathcal{N}_m(k)|$. Consider the optimization problem OPTSTAT(k), obtained by modifying the optimization problem OPTSTAT from Section 3.5.1 to account for the network state $\mathcal{N}(k)$, which is given below:

$$\begin{aligned}
\max_{\left((q_m(f))_{f\in\mathcal{F}}\right)_{m\in\mathcal{M}},\left((r_m(c))_{c\in\mathcal{C}_u}\right)_{m\in\mathcal{M}}} \quad & \sum_{m\in\mathcal{M}} N_m(k) U_m^E\Bigg(\mathbb{E}_m\left[q_m\left(F^\pi\right)\right] \\
& \qquad -U_m^V\left(\mathbb{E}_m\left[\left(q_m\left(F^\pi\right)-\mathbb{E}_m\left[q_m\left(F^\pi\right)\right]\right)^2\right]\right)\Bigg) \\
\text{subject to } \sum_{j\in\mathcal{N}(k)} c_j\left(r_{m(j)}\left(c_j\right)\right) &\leq \overline{c},\ \forall\ \left(c_j\right)_{j\in\mathcal{N}(k)} \in \mathcal{C}_u^{\mathcal{N}(k)}, \qquad (4.2)\\
q_m(f) &\geq 0,\ \forall\ f\in\mathcal{F},\ \forall\ m\in\mathcal{M}, \\
q_m(f) &\leq q_{\max},\ \forall\ f\in\mathcal{F},\ \forall\ m\in\mathcal{M}, \\
r_m(c) &\geq r_{i,\min},\ \forall\ c\in\mathcal{C}_u, \forall\ m\in\mathcal{M}, \\
p_m^d\frac{\mathbb{E}_m\left[F^\pi\left(q_m\left(F^\pi\right)\right)\right]}{\overline{p}_m} &\leq 1, \forall\ m\in\mathcal{M}, \\
\frac{\mathbb{E}_m\left[F^\pi\left(q_m\left(F^\pi\right)\right)\right]}{\left(1+\overline{\beta}_m\right)} &\leq \frac{\mathbb{E}_m\left[r_m\left(C^\pi\right)\right]}{\tau_{slot}}, \forall\ m\in\mathcal{M}.
\end{aligned}$$

where $\mathbb{E}_m[.]$ denotes the expectation with respect to the $\left(\pi_m^{\mathcal{F}}(f)\right)_{f\in\mathcal{F}}$ associated with class $m \in \mathcal{M}$. Thus, we have adapted the optimization problem OPTSTAT to obtain OPTSTAT(k) in which our goal is to find optimal quality choices $q_m(f)$ for $f \in \mathcal{F}$ and resource allocation to the video clients $r_m(c)$ for $c \in \mathcal{C}_u$, for each class m. Hence, $(r_m(c))_{c\in\mathcal{C}_u}$ is a class specific resource allocation policy, which prescribes a video client resource allocation of $r_m(c)$ to a video client with capacity function $c \in \mathcal{C}_u$ (in a given slot), and hence could result in different resource allocation in a slot to video clients of the same class with different capacity consumption functions. Similarly, we can view $(q_m(f))_{f\in\mathcal{F}}$ as a class specific quality adaptation policy.

One of the main goals of this section is to argue that NOVA has continuity properties with respect to the state of the network so that various quantities associated with NOVA do not change significantly for small changes in network state, and we establish this by studying continuity properties of an optimization problem OPTSTAT-HT(k) presented in the sequel which is obtained as an approximation of OPTSTAT(k). The continuity properties of OPTSTAT(k) are difficult to analyze especially due to (4.2) where the number of constraints depends exponentially on $\mathcal{N}(k)$. Hence, we study an approximation of OPTSTAT(k) for a heavy traffic regime. We consider a heavy traffic regime with network states $\mathcal{N}_m^{\Lambda}(k)$ where $\left|\mathcal{N}_m^{\Lambda}(k)\right| = \Lambda\left|\mathcal{N}_m(k)\right|$ for each $m \in \mathcal{M}$ so that the number of video clients in the network is scaled up by Λ, and we scale down the video client capacity consumption functions by Λ. That is, to accommodate the scaling up of video clients, we expand the capacity region by assuming that the video client capacity consumption functions take values in the set

$$\mathcal{C}_u^{\Lambda} = \left\{\frac{c}{\Lambda} : c \in \mathcal{C}_u\right\},$$

i.e., the set of video client capacity consumption functions obtained by scaling the video client capacity consumption functions in the original set $\mathcal{C}_u$. Also, let

$$\alpha_m(k,c) = \frac{\sum_{j \in \mathcal{N}_m^{\Lambda}(k)} I(c_{jk} = c)}{\sum_{m \in \mathcal{M}} |\mathcal{N}_m^{\Lambda}(k)|}, \tag{4.3}$$

denote the fraction of video clients in slot k that are of class m and have video client capacity consumption function $\frac{c}{\Lambda}$ for some $c \in \mathcal{C}_u$. Next, we make the following two assumptions:

Assumption-UI: For each class $m \in \mathcal{M}$, the evolution of video client capacity consumption functions of the video clients of class m are independent.

Assumption-CF: For each slot k, the ratio of total number of video clients to Λ is equal to $\overline{N}$, and the fraction of video clients of a class $m \in \mathcal{M}$ is equal to a fixed constant π_m.

The Assumption-UI is similar to that made in [32] which imposes a simple model on the correlation of the video client capacity consumption functions across video clients of the same class in any given slot. The Assumption-CF is essentially an assumption on the arrival process and holding times associated with the video clients of different classes. Although the assumption is very rigid and rarely holds in any practical system, the conditions in the assumption are roughly satisfied if the arrival processes are Poisson and holding times are exponentially distributed, and if the holding times are scaled up by Λ (so that the video clients stay in the network longer giving more time to 'learn' NOVA parameters associated with the stochastic network). Under Assumption-UI and CF, we see that $\alpha_m(k,c)$ converges to a constant $\pi_m \pi_m^{\mathcal{C}_u}(c)$ for almost all sample paths since the evolution of the video client capacity consumption functions is modeled as a stationary ergodic process. Motivated by this observation, we modify Constraint (4.2) in OPTSTAT(k) as given below:

$$N(k)\sum_{c\in\mathcal{C}_u}\sum_{m\in\mathcal{M}}\alpha_m(k,c)c\left(r_m\left(c\right)\right) \quad\leq\quad \bar{c}. \tag{4.4}$$

Hence, instead of requiring that constraints in (4.2) are met for each $(c_j)_{j\in\mathcal{N}^\Lambda(k)} \in \left(\mathcal{C}_u^\Lambda\right)^{\mathcal{N}^\Lambda(k)}$ (where $\left(\mathcal{C}_u^\Lambda\right)^{\mathcal{N}^\Lambda(k)}$ corresponds to the set of all possible combination of capacity consumptions function realizations for slot k given the network state $\mathcal{N}^\Lambda(k)$), we use the consequences of Assumptions UI and CF and only require that only an approximate version of (4.2) is satisfied.

We now define a new optimization problem OPTSTAT-HT(k) with objective

$$\max_{\left((q_m(f))_{f\in\mathcal{F}}\right)_{m\in\mathcal{M}},\left((r_m(c))_{c\in\mathcal{C}_u}\right)_{m\in\mathcal{M}}} \sum_{m\in\mathcal{M}}\sum_{c\in\mathcal{C}_u} N(k)\alpha_m(k,c)U_m^E\left(\mathbb{E}_m\left[q_m\left(F^\pi\right)\right]\right.$$
$$\left.-U_m^V\left(\mathbb{E}_m\left[\left(q_m\left(F^\pi\right)-\mathbb{E}_m\left[q_m\left(F^\pi\right)\right]\right)^2\right]\right)\right)$$

and with same constraints as OPTSTAT(k) except for (4.2) replaced by (4.4). Then, we have the following continuity result regarding the optimal solution and Lagrange multipliers associated with OPTSTAT-HT(k) with respect to the parameters $\left((\alpha'_m(k,c))_{c\in\mathcal{C}_u}\right)_{m\in\mathcal{M}}$ defined as $\alpha'_m(k,c){:=}N(k)\alpha_m(k,c)$ for each $c\in\mathcal{C}_u$ and $m\in\mathcal{M}$.

Lemma 4.1. *The optimal value of the objective of OPTSTAT-HT(k) is a continuous function of* $\left((\alpha'_m(k,c))_{c\in\mathcal{C}_u}\right)_{m\in\mathcal{M}}$.

Proof. The proof follows from Theorem 2.1 in [16]. □

Using Theorem 3.2, Corollary 3.1 and Lemma 3.2 (a) of Chapter 3, we can argue that the NOVA's performance is tied to the optimal value of OPTSTAT(k) and hence (roughly) to that of OPTSTAT-HT(k). Hence, using the above result, we can also argue that the limiting behavior of NOVA's performance is continuous in

$\left((\alpha'_m(k,c))_{c\in\mathcal{C}_u}\right)_{m\in\mathcal{M}}$. This observation along with the fact that $\alpha_m(k,c)$ is approximately $\pi_m\pi_m^{\mathcal{C}_u}(c)$ and $N(k)$ is approximately $\overline{N}$ for almost all slots, suggests that the video client dynamics in the network only causes small changes in the performance of NOVA, and that the performance is usually close to that the optimal value of OPTSTAT-HT(k) obtained with $\alpha'_m(k,c)$ set to $\overline{N}\pi_m\pi_m^{\mathcal{C}_u}(c)$.

4.6 Performance evaluation of NOVA via simulation

In this section, we carry out an evaluation of NOVA using Matlab simulations to compare the performance of a wireless network operating under NOVA vs one using Proportionally Fair (PF) network resource allocation (see [30]) and quality adaptation based on Rate Matching (RM). We discuss PF and RM in detail below. The main objective of this section is to use the simulation results to answer questions like:

- What are the typical gains under NOVA? In particular, we are interested in the following questions:
  - What are the typical capacity gains (defined later)?
  - How does NOVA perform in terms of rebuffering?
  - How do price constraints affect the gains?
  - Does NOVA penalize mean quality by too much to reduce variability in quality?
  - Is NOVA fair?
- What is the loss in the performance of NOVA (and for that matter, any good adaptation algorithm utilizing information about QR tradeoffs) in the absence of accurate QR tradeoffs? In particular, we are interested in the following questions:

– What is the reduction in performance if we have access only to partially accurate QR tradeoffs, e.g., those based on less sophisticated video quality assessment metrics like PSNR or if we just know 'averaged' QR tradeoffs?

– What is the reduction in performance if we do not have any information about QR tradeoffs?

4.6.1 Simulation setting

We consider a wireless network with $\tau_{slot} = 10$ msecs, and with allocation constraints of the form $c_k(\mathbf{r}_k) = \sum_{i \in \mathcal{N}} \frac{r_{i,k}}{p_{i,k}} - 1$ in each slot k, where $p_{i,k}$ denotes the peak resource allocation for video client i in slot k, i.e., if we only allocate resources to video client i in slot k, then $r_{i,k} = p_{i,k}$ is the maximum resource allocation to the video client that does not violate the allocation constraint in the slot. To obtain traces of peak resource allocation for the video clients, we generated 300 sequences of length 150000 each using Markov Chain Monte Carlo method, in such a way that the values in consecutive slots are positively correlated (the positive correlation reflects the correlation of the wireless channel in adjacent slots) and the marginal distribution of the stationary process is that of an appropriately scaled version of the sequence is equal to a distribution which is representative of capacities seen by a randomly placed wireless user with single antenna equalizer in an HSDPA system with 50% load (and thus associated interference) from its neighbors [1].

Unless mentioned otherwise, in our simulations, we consider settings with heterogeneous channels: we uniformly and at random pick a sequence for each video client from the 300 sequences, scale the sequence by a uniformly distributed random number in the range $[0.5, 1.5]$, and use the scaled sequence as the peak resource allocation seen by the video clients over 15000 slots. Thus, a video client with random scaling close to 0.5 sees the worst wireless channel on average, whereas one

[1] This data was provided by a service provider and is based on a simulation framework for such a system.

with random scaling close to 1.5 sees the best. We also present simulation results for a setting with homogeneous channels later in the section. In the setting with homogeneous channels, we uniformly at random pick a sequence for each video client from the 300 sequences and just use the sequence (without any additional scaling) as the peak resource allocation seen by the video clients.

Under PF, network resource allocation in slot k is an optimal solution to

$$\max_{\mathbf{r}} \left\{ \sum_{i \in \mathcal{N}} \frac{r_i}{\rho_{i,k}} : \; c_k(\mathbf{r}) \leq 0, \; r_i \geq r_{i,\min} \; \forall \; i \in \mathcal{N} \right\} \tag{4.5}$$

where the parameters $(\rho_{i,k})_{i \in \mathcal{N}}$ track the mean resource allocation to the video clients, and are updated using (3.73) with ϵ set to 0.01 (this is a good choice since $\rho_{i,k}$ is getting updated at a high rate of once every $\tau_{slots} = 10$ msecs).

In our simulations, we consider video clients downloading different parts of three open source movies Oceania (about 55 mins long), Route 66 (about 100 mins long) and Valkaama (about 90 mins long). The movies Oceania and Valkaama are compressed at rates 0.1, 0.2, 0.3, 0.6, 0.9 and 1.5 Mbps, with segments of duration 1 second each (hence, each segment is available in six representations). The movie Route 66 is compressed at rates 0.1, 0.2, 0.3, 0.6 and 0.9 Mbps, with segments of duration 1 second each (hence, each segment is available in five representations). Unless mentioned otherwise, in each simulation, video clients pick a movie and starting segment (index) for the movie at random, and start downloading the rest of the movie from that segment onwards. A video client on reaching the last segment of a movie continues viewing the movie from the first segment. We measure STQ of a representation using a proxy for DMOS (Differential Mean Opinion Score) score (see [36] for a discussion on DMOS) associated and mapping: STQ=100-DMOS. We chose this mapping as it roughly maps the proxy DMOS scores to the range $[0, 100]$, and an increase in STQ (unlike that for DMOS) corresponds to an improvement in quality. The proxy DMOS score for a representation is obtained from the value of the video quality assessment metric MSSSIM-Y (see [55]) associated with the representation, by using the following mapping (obtained from the model used in [46]): DMOS $= 13.6056 \times \log(1 + (1 - \text{MSSSIM-Y})/0.0006)$. The MSSIM-Y value of each segment was obtained as the average MSSIM-Y of the constituent frames obtained using the code given in [1]. To summarize, the QR tradeoffs used in our simulations map STQ values (obtained essentially from MS-SSIM metric) for five/six representations (each one second long) to its associated compression rate. Recall that this compression rate also accounts for the size of overheads due to metadata. The diversity of QR tradeoffs associated with these movies is illustrated in Figs. 4.3-4.5.

Till now, we assumed that we have a continuous set of quality choices for each segment. However, in practice, video segments are only available in a *finite* number of representations, and this is also the case with QR tradeoffs we obtained for the three movies. Thus for our simulations, we modify the optimization problem $\text{QNOVA}_i(\boldsymbol{\theta}_i, f_i)$, used in quality adaptation of NOVA, by imposing an additional restriction that the quality for segment s of video client i is picked from the finite set $\mathcal{Q}_{i,s}$ of available quality choices associated with the segment. We discuss this modification in more detail in Subsection 4.4.2 of Chapter 4.

In quality adaptation based on RM (Rate Matching), the video client essentially tries to 'match' the compression rate of the selected representation to (current estimate of) mean resource allocation in bits per second, and further modifies the selection to respond to the state of the playback buffer. This is basic feature in many compression rate adaptation algorithms, for instance, see [3] where (following their terminology) we see that 'requested bitrate' (i.e., size of the representation) stays close to the 'average throughput' (i.e., $\rho_{i,k}$ in our setting) in Microsoft Smooth Streaming player and Netflix player. For each video client $i \in \mathcal{N}$, the

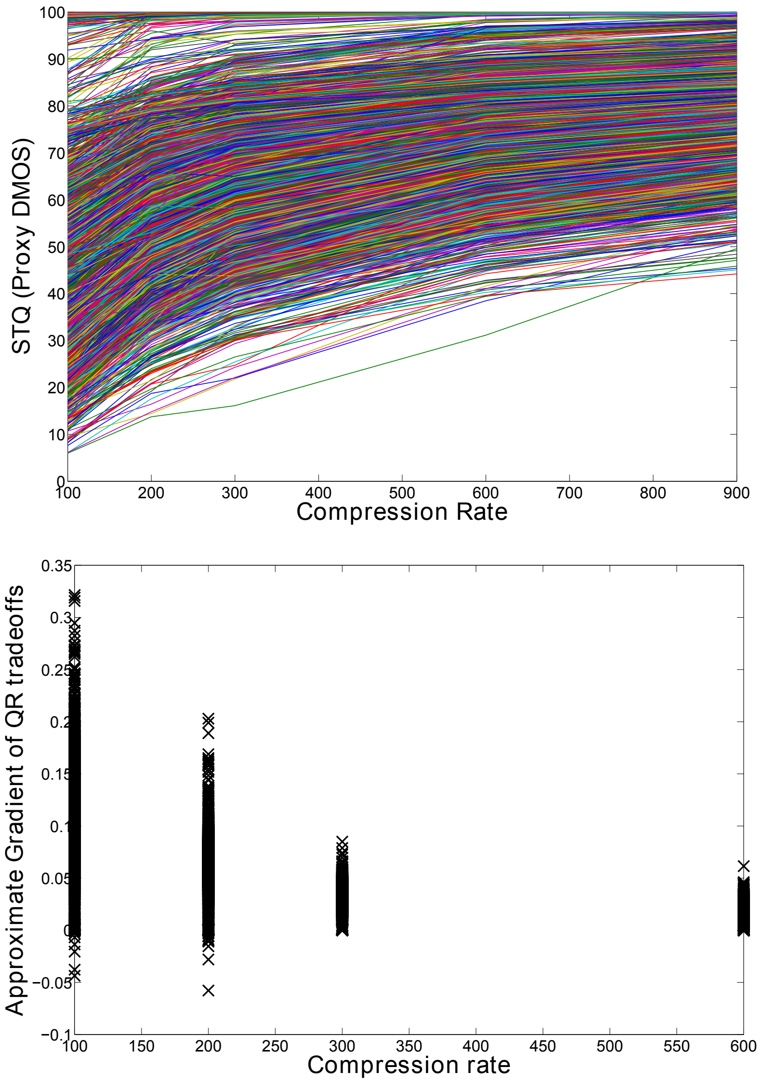

Figure 4.3: Diversity in QR tradeoffs of movie Route 66.

variables $I_i^{cautious}(k)$ and $I_i^{aggressive}(k)$ are used to enable RM to respond to low and high playback buffer respectively. The variable $I_i^{cautious}(k)$ is set to one if in slot k, the playback buffer has video content of duration less than 10 seconds and is set to zero if the playback buffer has video content of duration greater than 15 seconds. The variable $I_i^{aggressive}(k)$ is set to one if in slot k, the playback buffer

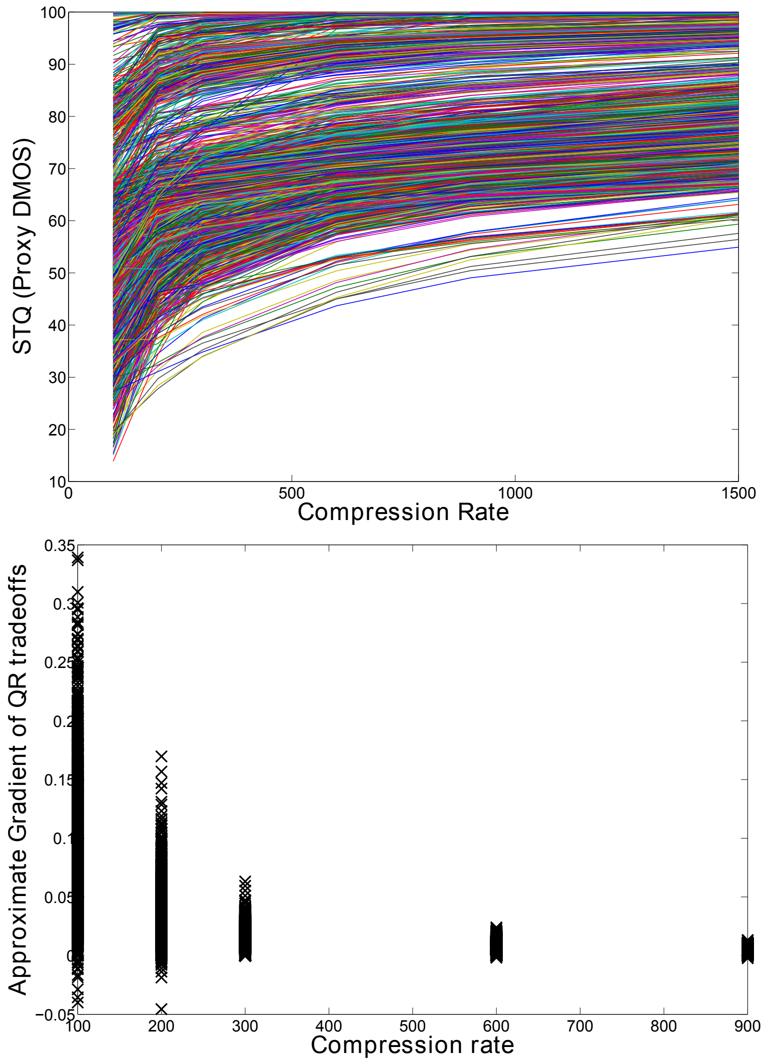

Figure 4.4: Diversity in QR tradeoffs of movie Oceania.

has video content of duration greater than 30 seconds and is set to zero if the playback buffer has video content of duration less than 25 seconds. The quality adaptation in RM works as follows: if any video client $i \in \mathcal{N}$ finishes download of $(s-1)$ th segment in slot k, we first find the representation with quality equal to $\text{argmax}_{q_i} \left\{ q_i \in \mathcal{Q}_{i,s} : \ f_{i,s}(q_i) \leq 0.99\rho_{i,k}, \ p_i^d f_{i,s}(q_i) \leq \overline{p}_i \right\}$, and let $M_{i,s}^{RM,0}$ denote the index of this representation. We denote the index of the representation picked by

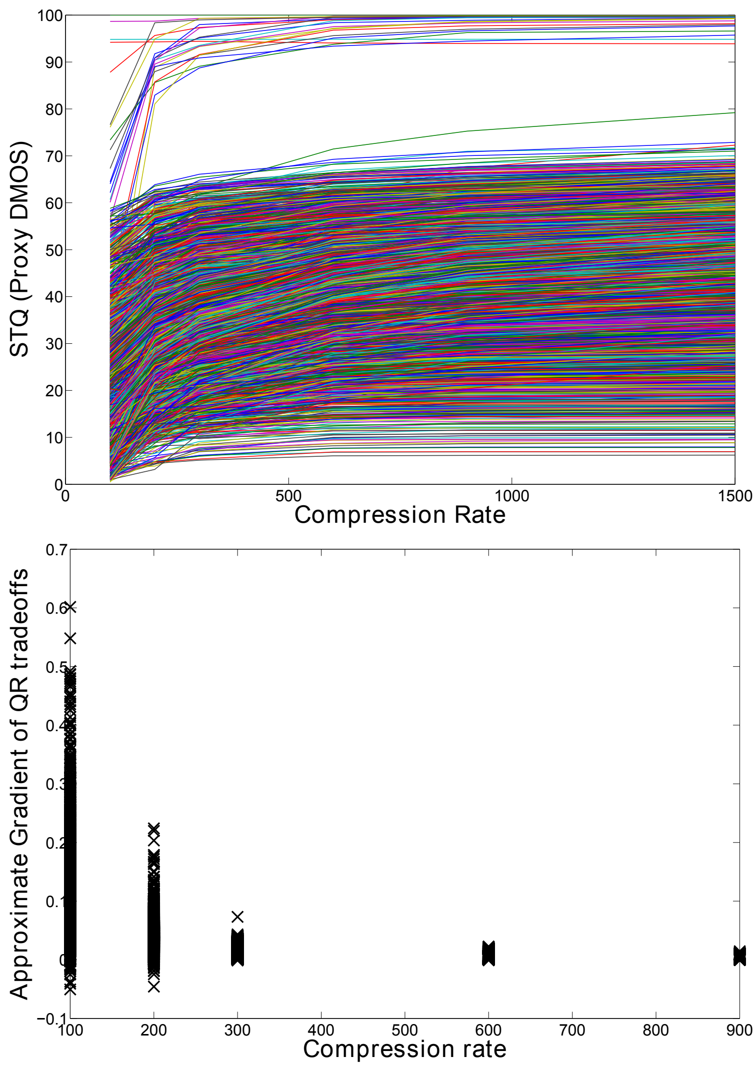

Figure 4.5: Diversity in QR tradeoffs of movie Valkaama.

RM for segment s of video client i by $M_{i,s}^{RM}$ which is given by

$$M_{i,s}^{RM} = \begin{cases} \max\left(\min\left(M_{i,s}^{RM,0} + I_i^{aggressive}(k) - I_i^{cautious}(k), M_{i,s,max}^{RM}\right), 1\right), \\ 1, \text{ if playback buffer has video content duration less than 5 secs,} \end{cases}$$

where $\overline{M}_{i,s}^{RM}$ is the number of representations available for segment s of video client

i. Hence, RM picks a lower representation if $I_i^{cautious}(k) = 1$ (i.e., when playback buffer is low), picks a higher representation if $I_i^{aggressive}(k) = 1$ (i.e., when playback buffer is high), and picks the lowest representation when the risk of rebuffering is high. Thus, RM meets the price constraint by ensuring that it is met for each segment. While considering price constraints in our simulations, we let $p_i^d = 0.01$ dollars per bit, and explicitly indicate the price constraints by referring to RM as RM($\overline{p}_i$) when we have cost constraint $\overline{p}_i$ for video client i.

.

For our simulations of NOVA, we let $\epsilon = 0.05$ and $r_{i,min} = 0.001$ bits. We set $U_i^E(e) = e_i$ and $U_i^V(v) = 0.05v_i$ for each $i \in \mathcal{N}$, and hence we only have to track the parameters $(m_{i,.}, b_{i,.}, d_{i,.})$ in this implementation of NOVA. We let $\overline{\beta}_i = 0$ for each $i \in \mathcal{N}$, and consider settings with two types of price constraints: in the first setting there are no price constraints, and in the second, each user $i \in \mathcal{N}$ has a price constraint of $\overline{p}_i = 3$ dollars per second. While evaluating the rebuffering time in the simulation results, we allow for a startup delay of 3 secs (which does not count towards rebuffering time). For each $i \in \mathcal{N}$, we chose $h_i^D(d_i) = 10d_i$ and

$$h_i^B(b_i) = h_{i,0}\left(\frac{b_i}{0.05} + \max\left(\frac{b_i - 20}{0.05}, 0\right)^2\right), \tag{4.6}$$

with $h_{i,0} = 0.005$ (see Subsection 4.4.2 for a discussion about this choice of $h_i^B(.)$). In all our simulations, we use the following initialization of NOVA parameters for each $i \in \mathcal{N}$: $m_{i,0} = 25$, $b_{i,0} = \frac{40}{0.05}$ and $d_{i,0} = 1$. Note that these initializations are used in all simulations ranging from lightly loaded (e.g. $N = 12$) to heavily loaded networks (e.g. $N = 33$), and for users seeing very good wireless channels to very bad wireless channels. Given the challenge of operating in these diverse settings, we enable NOVA to quickly 'learn' the setting by starting with larger values of ϵ for a few slots and segments initially and we keep reducing it until it reaches 0.1.

Each point in the plots discussed below is obtained by running the associated algorithm in 50 times where each simulation is run until all the users have downloaded a video of duration at least 10 minutes (i.e., 600 segments). Each point corresponds to a fixed number N of video clients in the network, and we vary N over the set $\{12, 15, 18, 21, 24, 27, 30, 33\}$. We refer to the combination of PF resource allocation and RM quality adaptation as PF-RM. To study the effectiveness of the quality adaptation in NOVA, we also study the performance of PF-QNOVA obtained by using PF resource allocation and the quality adaptation in NOVA. We refer to the modification of NOVA, PF-QNOVA and PF-RM with price constraint of 3 dollars per bit using the phrases NOVA(3), PF-QNOVA(3) and PF-RM(3) respectively. While implementing NOVA(3) and PF-QNOVA(3) with price constraint of 3 dollars per bit, we used a more stringent price constraint of 0.95×3 to ensure that the constraint is met for short videos (note that Theorem 3.1 guarantees that the constraint will be met for long enough videos without any additional tightening of the constraint).

4.6.2 Simulation results

In Fig. 4.6, we compare the QoE of the video clients under different algorithms, where we measure QoE using the metric QoE_1 which is the average across simulation runs of

$$\frac{1}{|\mathcal{N}|}\sum_{i\in\mathcal{N}}\left(m_i^{600}\left(q_i\right)-\sqrt{\mathrm{Var}_i^{600}\left(q_i\right)}\right),$$

where $m_i^{600}\left(q_i\right)-\sqrt{\mathrm{Var}_i^{600}\left(q_i\right)}$ is the metric proposed in [59] with the scaling constant for $\sqrt{\mathrm{Var}_i^{600}\left(q_i\right)}$ set to unity (and $m_i^{600}\left(q_i\right)$ and $\mathrm{Var}_i^{600}\left(q_i\right)$ are defined in (3.4) and (3.5)).

On comparing QoE_1 using Fig. 4.6, we see that NOVA performs much better

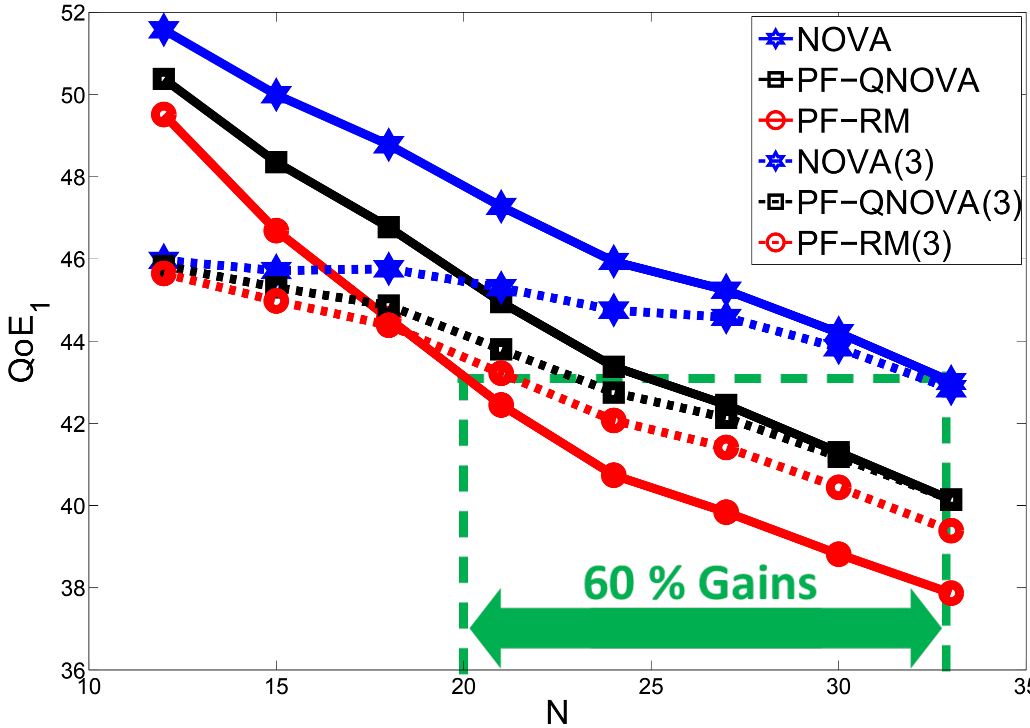

Figure 4.6: QoE_1 gains from NOVA.

than PF-RM and PF-QNOVA, and in fact provides 'network capacity gains' of about 60% over PF-RM, i.e., given a requirement on (user) average QoE_1, we can support about 60% more video clients by using NOVA than that under PF-RM. For instance, if we consider the horizontal dashed line in Fig. 4.6 that corresponds to an average QoE_1 requirement of about 43, we see that PF-RM can only support 20 video clients while meeting this requirement whereas NOVA can support almost 33 video clients. Under price constraint (of 3 dollars per second) also, we see that NOVA(3) provides network capacity gains of about 60% over PF-RM(3).

The gain from the adaptation component of NOVA is also visible in Fig. 4.6, where we see that PF-QNOVA provides network capacity gains of about 25% over PF-RM respectively.

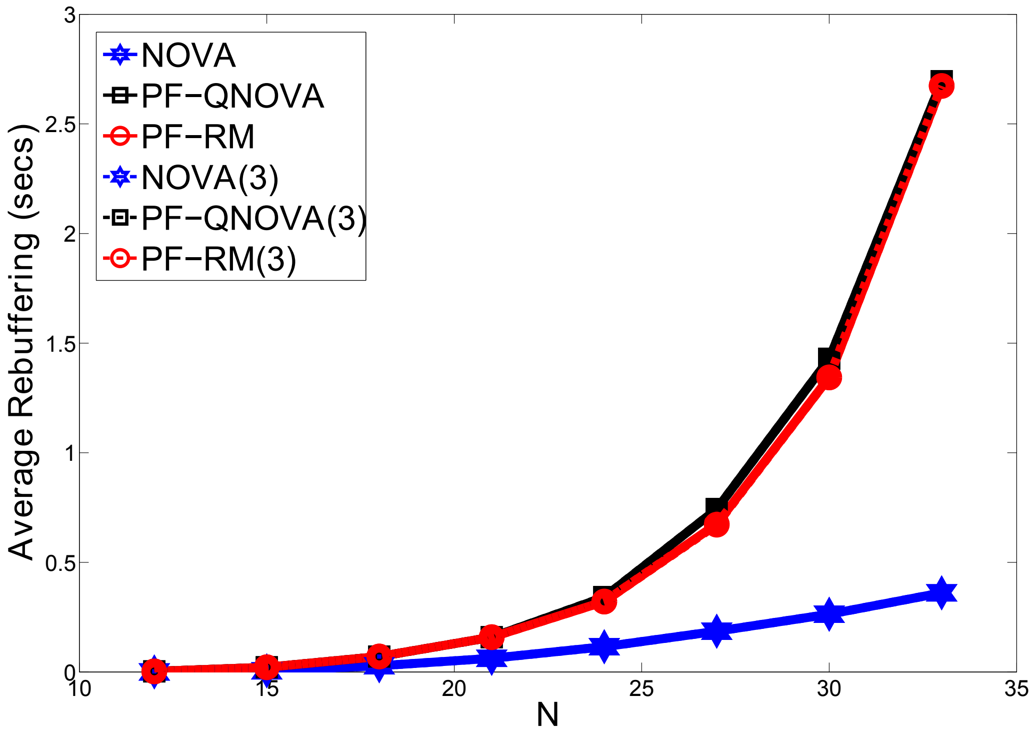

Figure 4.7: Reduction in rebuffering using NOVA.

The results in Fig. 4.7 depict the significant reduction in the amount of time spent rebuffering under NOVA and NOVA(3). Using Figs 4.6-4.7, we see that NOVA outperforms PF-RM in both the metric QoE_1 and the amount of time spent rebuffering which cover some of the most important factors affecting users' QoE (see the discussion in Section 3.1).

In Fig. 4.8, we compare the performance of different algorithms using another metric QoE_2 which is the average across simulation runs of

$$\frac{1}{|\mathcal{N}|}\sum_{i\in\mathcal{N}}\left(m_i^{600}\left(q_i\right)-\sqrt{\text{MSD}_i^{600}\left(q_i\right)}\right),$$

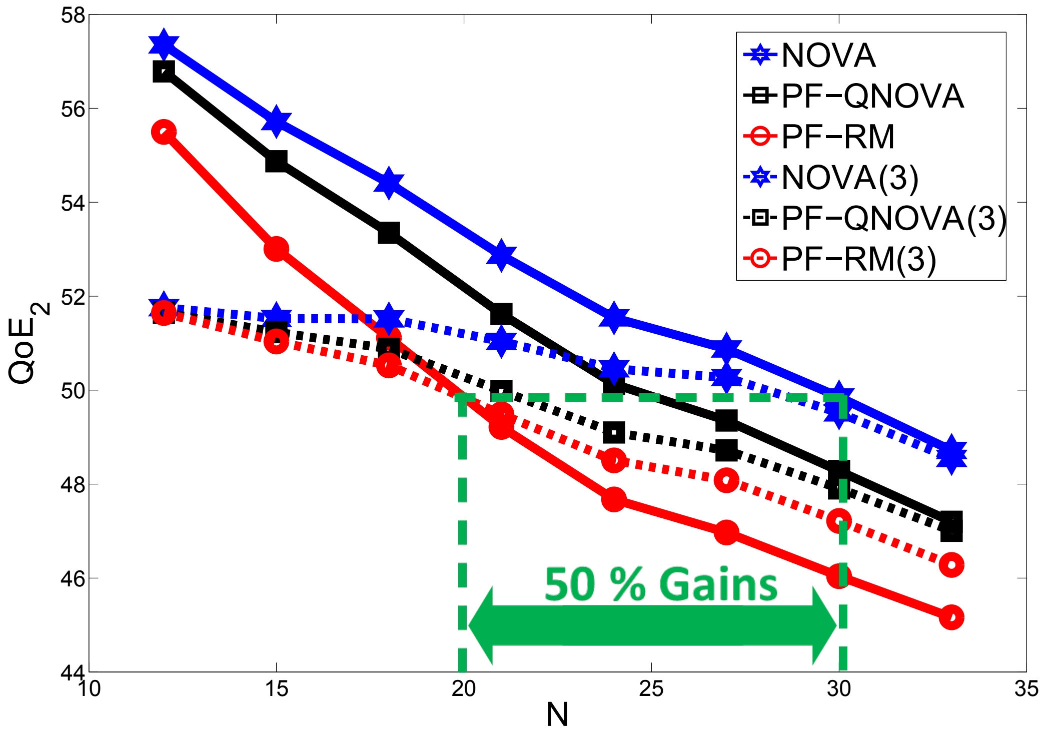

Figure 4.8: QoE_2 gains from NOVA.

where

$$\text{MSD}_i^{600}(q_i) := \frac{1}{600}\sum_{s=1}^{600}(q_{i,s+1} - q_{i,s})^2 .$$

Note that this metric is similar to the metric QoE_1, except that it penalizes short term variability in quality (i.e., variability across consecutive segments). From Fig. 4.8, we see that NOVA provides gains similar to those in the case of the metric QoE_1 (in Fig. 4.6), By comparing QoE_2, we see that NOVA and NOVA(3) provide network capacity gains of about 50% over PF-RM and PF-RM(3) respectively.

The results in Fig. 4.9 show that the improvement in QoE_1 and QoE_2 under NOVA does not come at the cost of significant reduction in mean quality. In fact, the results suggest that NOVA has better mean quality (in addition to lower variability in quality) in all but lightly loaded networks (i.e., $N = 12$). Also note that we can

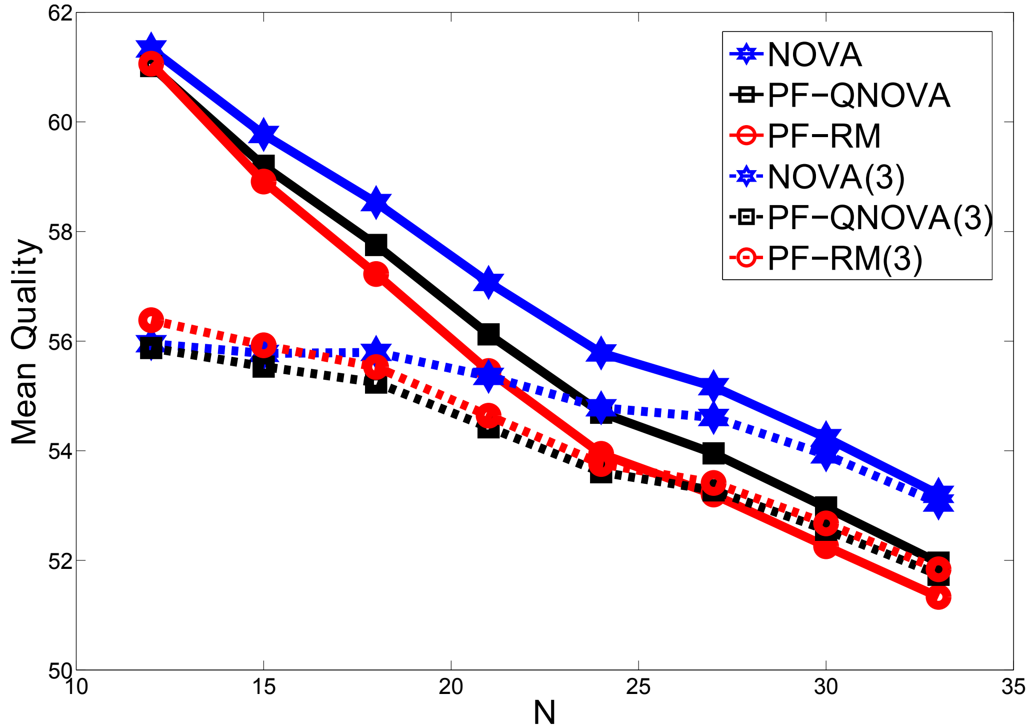

Figure 4.9: Mean quality gains from NOVA.

further increase the mean quality under NOVA (at the cost higher variability in quality) if we scale down the functions $\left(U_i^V\right)_{i\in\mathcal{N}}$.

The results in Fig. 4.10 indicate that, when compared to PF-RM, NOVA is more fair in QoE_1 delivered to the video clients. Here, we measure fairness as (the average across simulations of) the ratio of the difference between maximum and minimum of QoE_1 across users to the mean (across users of) QoE_1. Although we chose $\left(U_i^E\right)_{i\in\mathcal{N}}$ to be linear functions, the fairness associated with NOVA in these results can be attributed to the concavity of inverse of QR tradeoffs (i.e., convexity of QR tradeoffs) and the structure of the objective function (see (3.18)) of the optimization problem associated with quality adaptation. Further, from Fig. 4.11, we see that NOVA(3) meets cost constraints (of 3 dollars per second).

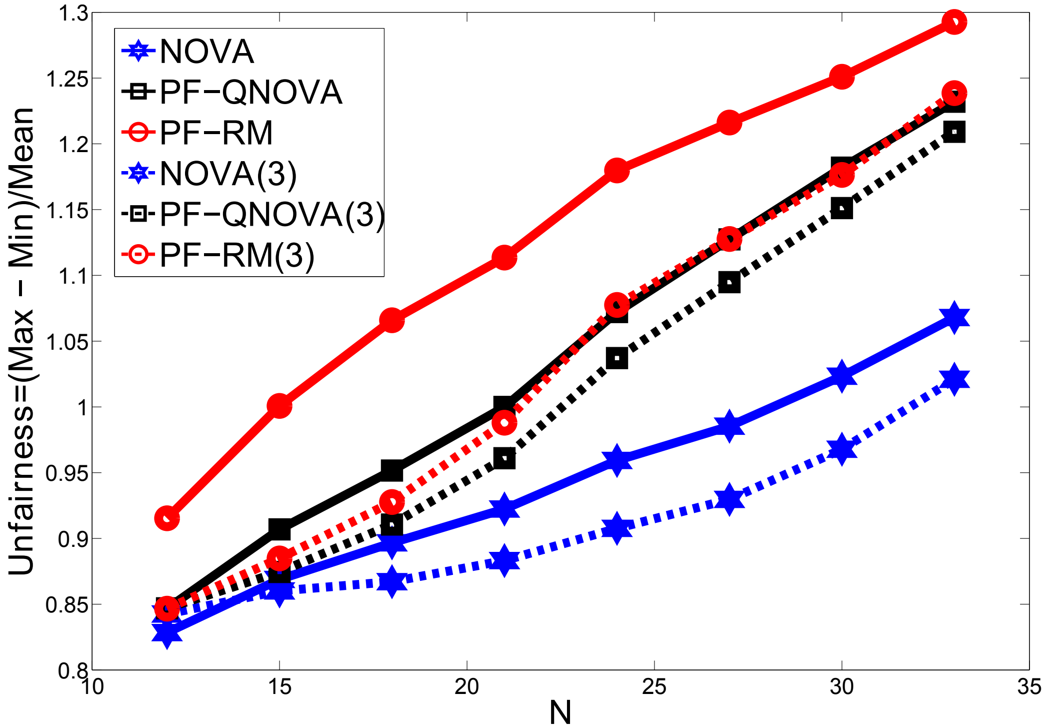

Figure 4.10: Fairness gains using NOVA.

We have depicted the results obtained using simulations for the setting with homogeneous channels (see the discussion about the setting in the beginning of this section) in Fig. 4.12. We see that the performance gains under homogeneous channels are slightly higher than those in the case of heteregeneous ones.

To assess the value of knowing accurate QR tradeoffs, we carried out simulations for NOVA with STQ based on less sophisticated video quality assessment metrics. In particular, we carried out simulations where we used the same QR tradeoff for all segments and this QR tradeoff was equal to the average of QR tradeoffs of all segments of the movie being viewed by the video client. Thus, instead of using segment level QR tradeoff information, we are using an approximation based on long term features of the videos being viewed by the video clients. The results asso-

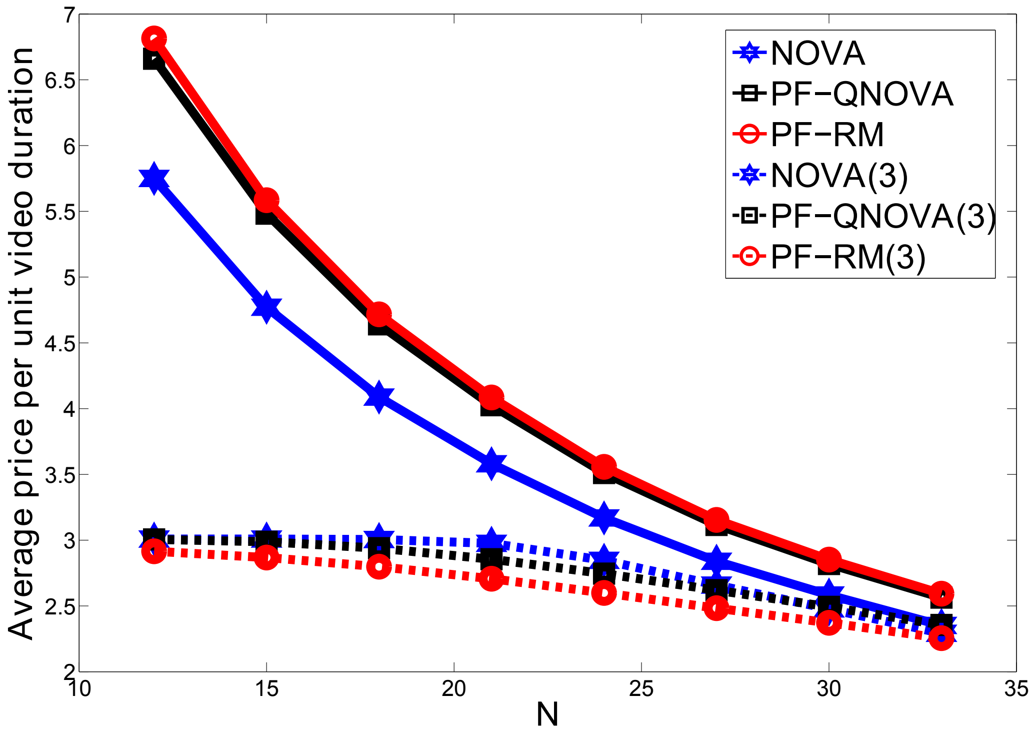

Figure 4.11: NOVA meets cost constraints.

ciated with this setting is depicted using the curve NOVA-Avg-QR in Fig. 4.13. We also carried out simulations using NOVA with STQ equal to PSNR, and the results associated with this setting are depicted using the curve NOVA-PSNR in Fig. 4.13. We picked $h_{i,0}$ in (4.6) as 0.0025 for the simulations with STQ equal to PSNR. In a setting, where we do not have any information about the QR tradeoffs, we could use crude metrics like $10\log(\text{Representation Size})$, and the results associated with this setting are depicted using the curve NOVA-No-QR in Fig. 4.13. Although NOVA-No-QR has (expectedly) the worst performance, we see that there is a significant reduction in performance (i.e., QoE_1) when we do not have accurate segment level QR tradeoff information, and further these reductions are not too different for NOVA-Avg-QR or NOVA-PSNR when compared to that for NOVA-No-QR.

We also compared the performance of NOVA, PF-QNOVA and PF-RM for the movies Oceania, Route 66 and Valkaama separately, and these results are depicted in Figs. 4.14-4.16.

4.7 Implementing NOVA: An example

In this section, we discuss an example for an implementation of NOVA for a network shared by video clients and data users. To simplify the exposition, we will make simplifying assumptions. We also discuss the issues related to signaling requirements, information exchange and complexity towards the end.

4.7.1 Setting

We consider a base station in a cellular network supporting a dynamic number of video clients and data users, i.e., the base station is the network controller responsible for network resource allocation. Let $\mathcal{N}(k)$ denote the set of video clients and $\mathcal{N}_D(k)$ denote the set of data users in the network in slot k. The priority given to video clients is determined by the parameter $p_V > 0$ (discussed in Section 4.3). Let the duration of each slot be equal to τ_{slot} seconds and that of each segment be l_{seg} seconds.

In each slot k, let $c_k\left((r_{i,k})_{i\in\mathcal{N}(k)\cup\mathcal{N}_D(k)}\right) \leq 0$ defined below

$$c_k\left((r_{i,k})_{i\in\mathcal{N}(k)\cup\mathcal{N}_D(k)}\right) = \sum_{i\in\mathcal{N}(k)} \frac{r_i}{p_{i,k}} + \sum_{i\in\mathcal{N}_D(k)} \frac{r_i}{p_{i,k}} - 1, \tag{4.7}$$

describe the capacity region in slot k, where $p_{i,k}$ and $p_{j,k}$ denotes the peak rate seen by video client $i \in \mathcal{N}(k)$ and data user $j \in \mathcal{N}_D(k)$ respectively in slot k, i.e., $p_{i,k}$ is the maximum rate that can be allocated to video client $i \in \mathcal{N}(k)$ in slot k when we allocate all the resources to this video client (and none to others). We assume that the video clients have no cost constraints (so that we can ignore the variables $d_{i,s}$) and that the QoE model is given by

$$e_i^S(q_i) = m_i^S(q_i) - c_v \mathrm{Var}^S(q_i),$$

for some positive constant c_v. Further, we set $U_i^E(e) = e$ for each $i \in \mathcal{N}(k)$. Also, let $\overline{\beta}_i = 0$ for each $i \in \mathcal{N}(k)$. We also set $r_{i,min} = 0$ for each $i \in \mathcal{N}$ (and ignore the requirement that it should be positive). For each data user $j \in \mathcal{N}_D(k)$, we use the following QoE model (see Section 4.3 for a discussion of this QoE model)

$$e_j^{K_S,data}(r_j) = m_j^{K_S}(r_j).$$

For the sth segment of video client $i \in \mathcal{N}(k)$, $\mathcal{Q}_{i,s}$ is the (finite) set of available quality choices for the segment and $f_{i,s}(q)$ denotes the compression rate of the representation associated with a quality choice $q_i \in \mathcal{Q}_{i,s}$ (see Subsection 4.4.2 for a detailed discussion).

4.7.2 Detailed algorithm

As for the proportional fair scheduler, the base station uses $\rho_{j,k}$ to track the mean rate allocation to data user $j \in \mathcal{N}_D(k)$. Since the update of variable $b_{i,k}$ requires the knowledge of segment download completions (see the update rule (3.24) of NOVA), the base station either has to be able look at the data stream of video clients to infer segment download completions, or rely on signaling from the video clients that indicate segment download completions. We focus on the latter setting in this section. Let the base station store an current estimate $b_{i,k}^B$ of $b_{i,k}$ for each video client $i \in \mathcal{N}(k)$, and we discuss the update rule for this estimate in the algorithm presented later in the section. Each video client $i \in \mathcal{N}(k)$ uses index s_i to track the number of segments downloaded by video client $i \in \mathcal{N}(k)$, the parameter $m_{i,s}$ to track mean quality, and uses the parameter $b_{i,k}$ to obtain a proxy for the risk of rebuffering. Thus, each video client $i \in \mathcal{N}(k)$ stores the current value of s_i, $m_{i,\cdot}$ and $b_{i,\cdot}$. The base station stores the current value of variable $\rho_{j,k}$ for each data user $j \in \mathcal{N}_D(k)$ in addition to the current value of variable $b^B_{i,k}$ for each video client $i \in \mathcal{N}(k)$. The detailed algorithm is given below.

INIT: The base station initializes $b^B_{i,0}$ for each video client $i \in \mathcal{N}(k)$, and $\rho_{j,0}$ for each data user $j \in \mathcal{N}_D(k)$. Each video client $i \in \mathcal{N}(k)$ initializes $m_{i,0}$ and $b_{i,0}$. Let $s_{i,0} = 1$ for each video client $i \in \mathcal{N}(k)$.

In each slot $k \geq 0$, carry out the following steps:

RNOVA-BS: At the beginning of slot k, base station *estimates* current capacity region c_k, and allocates rate $r^*_{i,k}$ to each video client $i \in \mathcal{N}(k)$ and $r^*_{j,k}$ to each data user $j \in \mathcal{N}_D(k)$, where $\left(r^*_{i,k}\right)_{i \in \mathcal{N}(k) \cup \mathcal{N}_D(k)}$ is an optimal solution to

$$\max_{\mathbf{r} \geq 0} \quad p_V \sum_{i \in \mathcal{N}(k)} h(b^B_{i,k}) r_i + \sum_{j \in \mathcal{N}_D(k)} \frac{r_j}{\rho_{j,k}} \tag{4.8}$$

$$\text{subject to} \quad c_k \left(\left(r_{i,k}\right)_{i \in \mathcal{N}(k) \cup \mathcal{N}_D(k)} \right) \leq 0, \tag{4.9}$$

and update

$$\begin{aligned} b^B_{i,k+1} &= b^B_{i,k} + \epsilon \left(\tau_{slot}\right), \ \forall \ i \in \mathcal{N}(k) \\ \rho_{j,k+1} &= \rho_{j,k} + \epsilon \left(r^*_{j,k} - \rho_{j,k}\right), \ j \in \mathcal{N}_D(k). \end{aligned}$$

RNOVA-VC: At the beginning of each slot k, each video client $i \in \mathcal{N}(k)$ updates

$b_{i,k}$ as follows

$$b_{i,k+1} = b_{i,k} + \epsilon\left(\tau_{slot}\right), \tag{4.10}$$

QNOVA-VC: During slot k, if any video client $i \in \mathcal{N}(k)_v$ finishes transmission of s_i th segment, the video client i sends an END-OF-SEG$_i$ to the BS, retrieves (this should be locally available at this time) $\mathcal{Q}_{i,s_i+1}$ and f_{i,s_i+1} for the $(s+1)$th (i.e., next) segment, and picks segment corresponding to quality q^*_{i,s_i+1} obtained using

$$q^*_{i,s_i+1} = \text{argmax}_{q_i \in \mathcal{Q}_{i,s_i+1}} \left(q_i - c_v \left(q_i - m_{i,s_i}\right)^2 - h(b_{i,k}) f_{i,s_i+1}\left(q_i\right) \right) \tag{4.11}$$

and update m_{i,s_i+1}, $b_{i,k+1}$ and s_i as follows:

$$m_{i,s_i+1} = m_{i,s_i} + \epsilon\left(q^*_{i,s_i+1} - m_{i,s_i}\right), \tag{4.12}$$

$$b_{i,k+1} = b_{i,k+1} - \epsilon\left(l_{seg}\right), \tag{4.13}$$

$$s_i = s_i + 1.$$

BS-REC-SIG: On receiving END-OF-SEG$_i$ from video client $i \in \mathcal{N}(k)$, the base station updates (overwrites)

$$b^B_{i,k+1} = b^B_{i,k+1} - \epsilon\left(l_{seg}\right), \tag{4.14}$$

Signaling required between the base station and video clients

Each video client $i \in \mathcal{N}(k)$ sends an END-OF-SEG$_i$ signal to the base station when it completes downloading a segment. Here, END-OF-SEG$_i$ can be viewed as a control signal carrying ID of the video client sending it and data to indicate that this is an END-OF-SEG control signal. When the base station receives the END-OF-SEG$_i$ signal from video client $i \in \mathcal{N}(k)$, it updates $b^B_{i,k}$ using (4.14). It can be verified that the update mechanism along with the signaling ensures that $b^B_{i,k}$, which is base stations's estimate for $b_{i,k}$, is equal to $b_{i,k}$ most of the time (except for the time duration between sending END-OF-SEG$_i$ and the reception of the signal at the basestaion). Note that if the segment lengths are variable, the video clients will also have to send the length of the next segment being downloaded as it is required (see 3.24) by the base station to update $b^B_{i,k+1}$.

Information flow

NOVA uses two types of 'external' data- (a) channel capacity data $c_.$, and (b) $\mathcal{Q}_{i,.}$ and $f_{i,.}$ for each video client $i \in \mathcal{N}(k)$. They are described in more detail below.

We assume that the base station knows (or knows with reasonable accuracy) the current value of c_k for each slot k, e.g., the current value of peak rates $\left((p_{i,k})_{i \in \mathcal{N}(k)}, (p_{j,k})_{j \in \mathcal{N}_D(k)} \right)$ for each video client in each slot k. The video clients could measure this and inform the base station.

For each video client $i \in \mathcal{N}(k)$, on completion of download of segment s, we assume that video client knows $\mathcal{Q}_{i,s+1}$ and $f_{i,s+1}(.)$ for the $s+1$th segment. One simple way to ensure this is to make sure that when a video client starts downloading a certain segment, the video client has requested the video server to ensure that video server has sent $\mathcal{Q}_{i,.}$ and $f_{i,.}(.)$ for the next few segments to the video client. Note that this is not a difficult requirement to meet even for live videos (and clearly not for stored videos). In the worst case, if this information is not available, then we could use a concave function (e.g., $\log(.)$) of the size of the segment as a proxy for quality.

The flow of information across various layers of the network for this implementation of NOVA is depicted in Fig. 4.17.

Complexity

The optimization problem used in resource allocation (described in (4.8)- (4.9)) is a linear program. Further, we can exploit the structure of allocation constraints given in (4.7) and linearity of objective, to show that will be carrying out optimal resource allocation if we pick a video client or data user that has the highest value a metric, and assign the peak rate to that video client or data user. The metric for this setting is equal to $p_V h(b^B_{i,k}) p_{i,k}$ for video clients $i \in \mathcal{N}(k)$, and $\frac{p_{j,k}}{\rho_{j,k}}$ for data users $j \in \mathcal{N}_D(k))$.

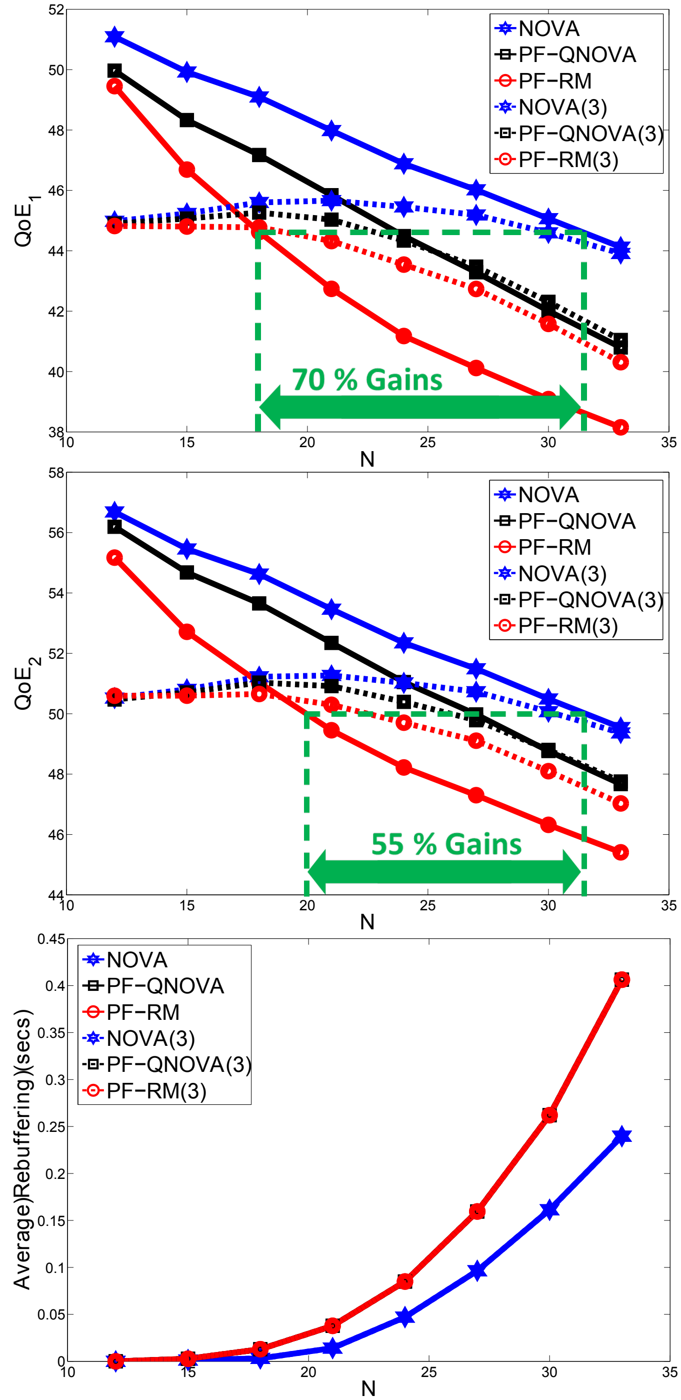

Figure 4.12: Performance gains using NOVA: Homogeneous channels.

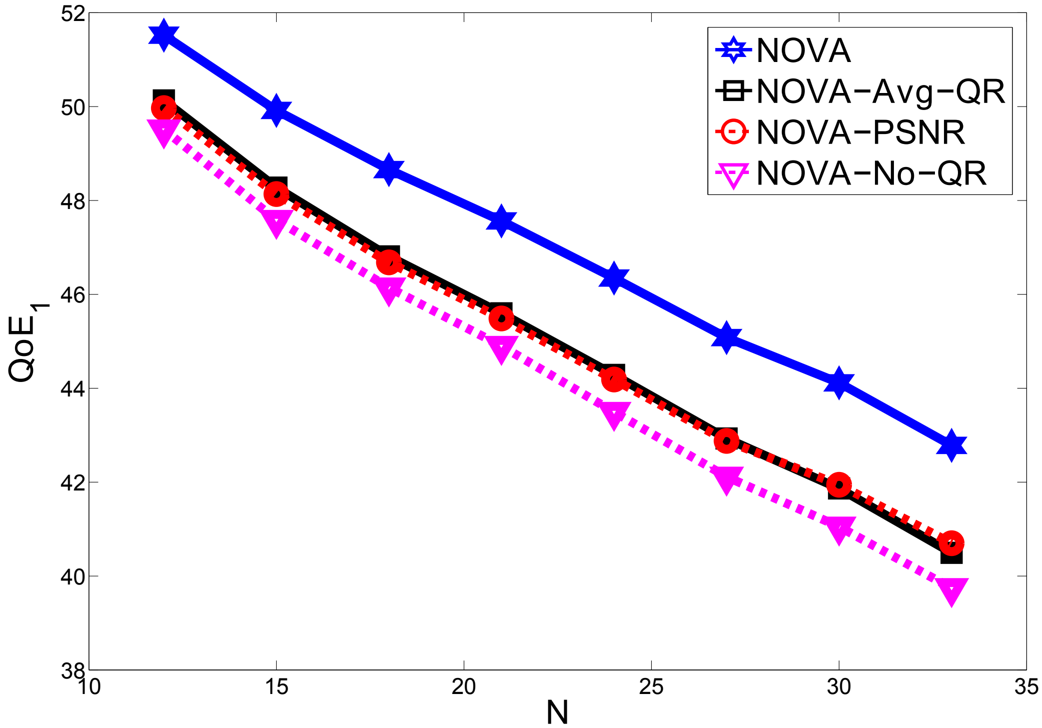

Figure 4.13: Value of knowing QR tradeoffs.

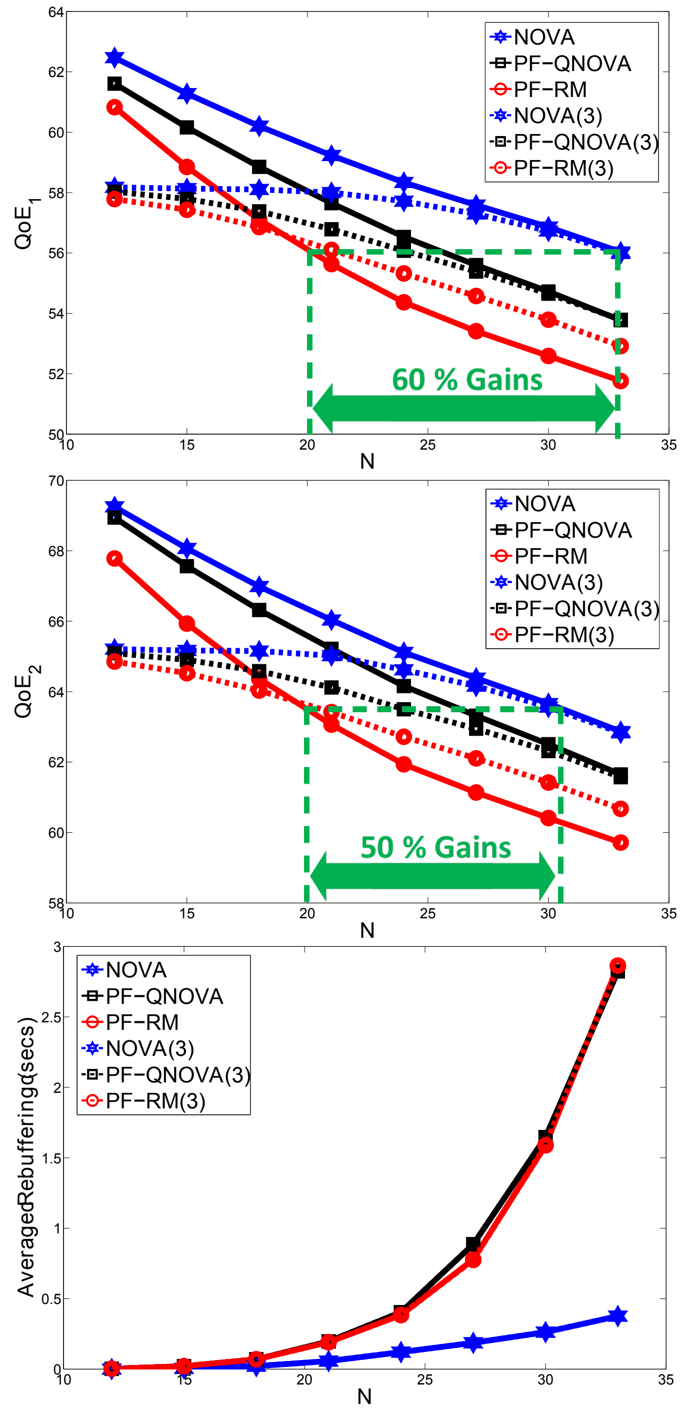

Figure 4.14: Performance gains using NOVA: Streaming movie Oceania.

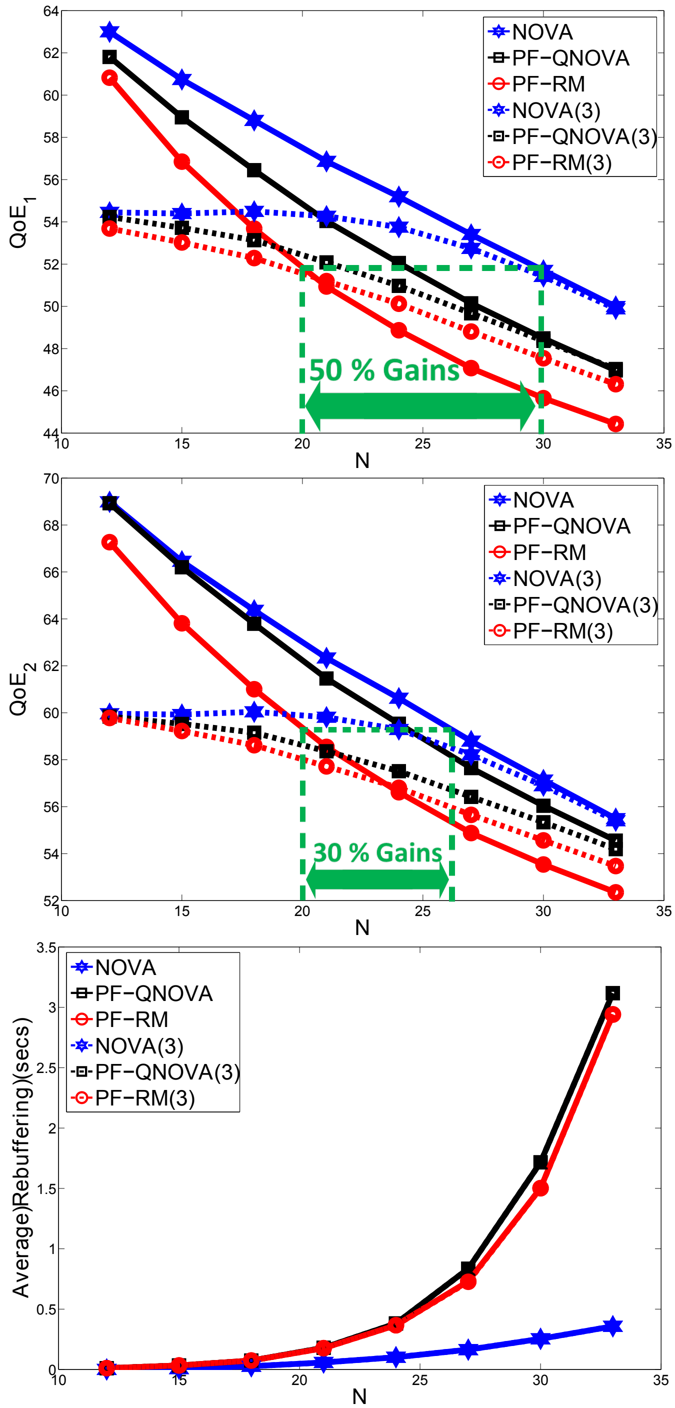

Figure 4.15: Performance gains using NOVA: Streaming movie Route 66.

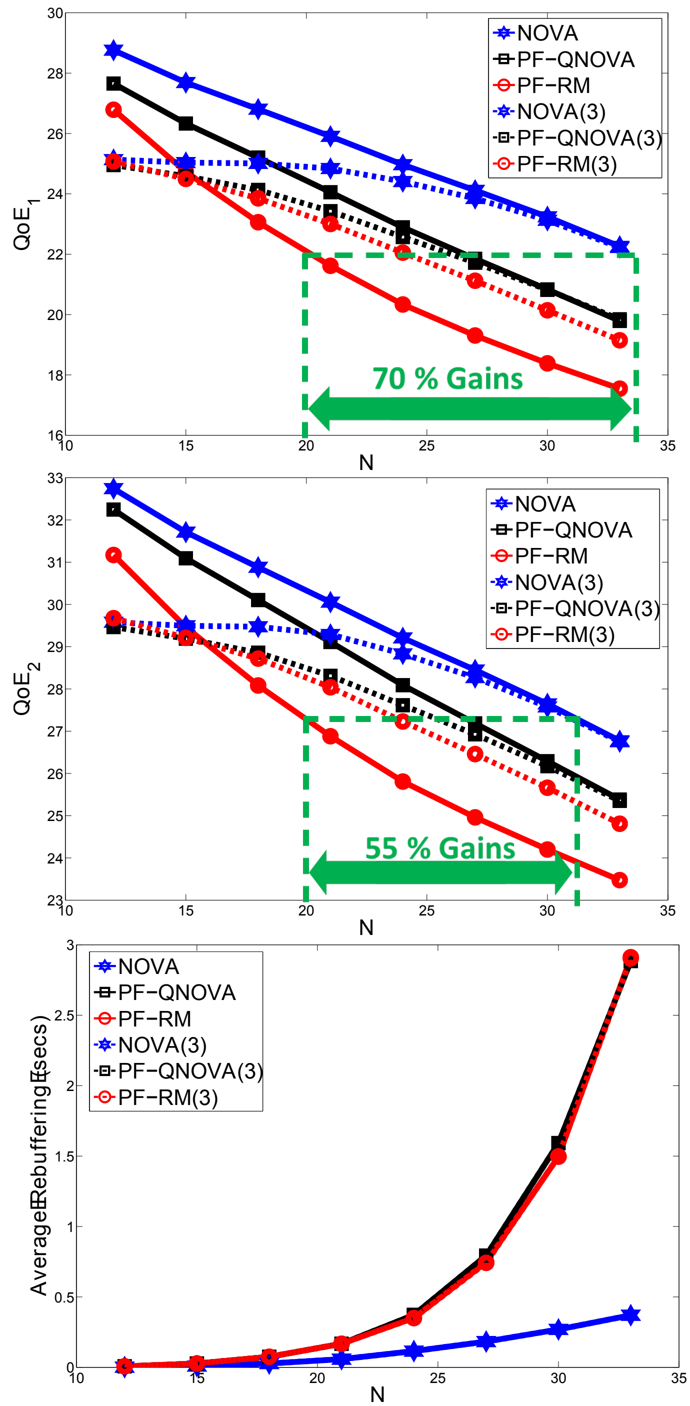

Figure 4.16: Performance gains using NOVA: Streaming movie Valkaama.

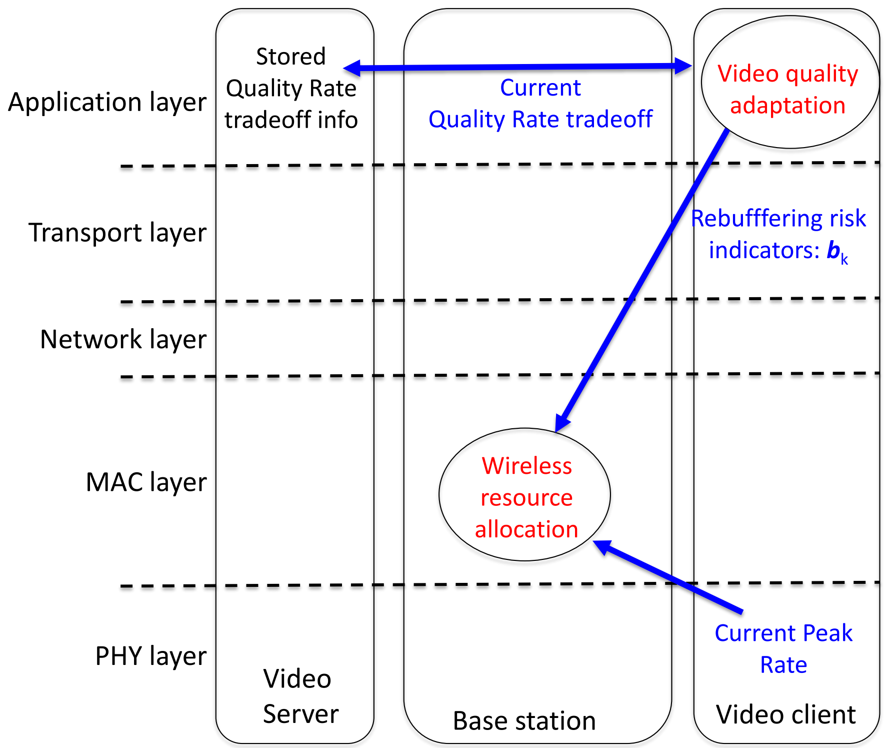

Figure 4.17: Cross Layer Information Flow

Chapter 5

Future Directions

5.1 A general approach for classes of online stochastic optimization problems

The work in this thesis suggests perhaps the possibility of a more general framework for certain classes of online stochastic optimization problems. By identifying the key properties that allowed the design of simple online algorithms for the problem considered in Chapter 2, we can obtain insights into such generalizations. These properties are not restricted to the specific application (i.e., realizing Mean-Variability-Fairness Tradeoffs in network resource allocation) considered in Chapter 2, and we are exploring new applications for using the general framework.

The algorithms presented in this thesis are computationally lightweight, and thus, can also be used in solving large offline stochastic optimization problems (such as OPT(T) considered in Chapter 2). We are also exploring the potential of our algorithms (and their generalizations) as attractive alternatives for solving similar large offline optimization problems.

5.2 Extensions to optimization in stochastic networks

In this thesis, we mainly focused on developing algorithms (with optimality guarantees) for network settings involving a fixed number of users. However, many real world networks are stochastic networks, i.e., there will be a time varying number of users utilizing network resources due to arrival and departure of users. Furthermore, solutions obtained for a network with fixed number of users need not always be optimal for stochastic networks, for e.g., [54, 39] point out that algorithms that are throughput optimal for a network with a fixed number of users need not be so in stochastic settings. Although we were able to study the performance of NOVA under simplifying assumptions in Chapter 4, the problem of designing *optimal* algorithms for stochastic networks to realize Mean-Variability-Fairness Tradeoffs is still open.

5.3 Rate of convergence

In this thesis, we developed simple online algorithms which have strong optimality properties under an appropriate convergence behavior. In our work, the theoretical analyses were focused primarily on establishing convergence, and we studied issues related to rate of convergence only using simulations. An important future direction of work would be to explore the possibility of providing guarantees on rate of convergence and how it is impacted by system characteristics, e.g., heterogeneity. This in turn will provide an idea of the time required by these algorithms to 'learn' its system and perform in an optimal manner. Further, ideas similar to those used in Newton method (see Chapter 9 in [10]) and [58], that use second order properties of the optimization problem may also be useful in developing algorithms with better convergence rates.

Bibliography

[1] LIVE software releases. http://live.ece.utexas.edu/research/Quality/index.htm.

[2] Rajeev Agrawal and Vijay Subramanian. Optimality of certain channel aware scheduling policies. *Allerton Conference on Communication, Control and Computing*, 2002.

[3] Saamer Akhshabi, Sethumadhavan Narayanaswamy, Ali C. Begen, and Constantine Dovrolis. An experimental evaluation of rate-adaptive video players over HTTP. *Signal Processing: Image Communication*, 2012.

[4] Dimitri Bertsekas. *Dynamic programming and optimal control (2 Vol Set)*. Athena Scientific, Jan. 2007.

[5] Dilip Bethanabhotla, Giuseppe Caire, and Michael J. Neely. Joint transmission scheduling and congestion control for adaptive video streaming in small-cell networks. *arXiv:1304.8083 [cs.NI]*, 2013.

[6] J. Frederic Bonnans and Alexander D. Ioffe. Quadratic growth and stability in convex programming problems with multiple solutions. *Journal of Convex Analysis*, 1995.

[7] J. Frederic Bonnans and Alexander Shapiro. Optimization problems with perturbations, a guided tour. *SIAM REVIEW*, 1996.

[8] J. Frederic Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer, 2000.

[9] Anna Bouch and M. Angela Sasse. Network quality of service: What do users need? *International Distributed Conference*, 1999.

[10] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2009.

[11] Yan Cai, Tilman Wolf, and Weibo Gong. Delaying transmissions in data communication networks to improve transport-layer performance. *IEEE Journal on Selected Areas in Communications*, May 2011.

[12] CISCO. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010-2015. Feb. 2011.

[13] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2006.

[14] Philippe de Cuetos and Keith W. Ross. Adaptive rate control for streaming stored fine-grained scalable video. *NOSSDAV*, 2002.

[15] Chao Chen et al. Rate adaptation and admission control for video transmission with subjective quality constraints. *In preparation.*

[16] Anthony V. Fiacco and Yo Ishizuka. Sensitivity and stability analysis for nonlinear programming. *Annals of Operations Research*, 1990.

[17] A.V. Fiacco and J. Kyparisis. Convexity and concavity properties of the optimal value function in parametric nonlinear programming. *Journal of Optimization Theory and Applications*, 1986.

[18] Jerzy A. Filar, L. C. M. Kallenberg, and Huey M. Lee. Variance-penalized Markov Decision Processes. *Mathematics of Operations Research*, 1989.

[19] Frances X. Frei, Ravi Kalakota, Andrew J. Leone, and Leslie M. Marx. Process variation as a determinant of bank performance: Evidence from the retail banking study. *Management Science*, 1999.

[20] Fangwen Fu and Mihaela van der Schaar. A systematic framework for dynamically optimizing multi-user wireless video transmission. *IEEE Journal on Selected Areas in Communications*, Apr. 2010.

[21] Robert M. Gray. *Probability, Random Processes, and Ergodic Properties.* Springer Publishing Company, 2nd edition, 2009.

[22] Bruce Hajek. *An Exploration of Random Processes for Engineers.* Jul. 2011.

[23] Jianwei Huang, Zhu Li, Mung Chiang, and A.K. Katsaggelos. Joint source adaptation and resource allocation for multi-user wireless video streaming. *IEEE Transactions on Circuits and Systems for Video Technology*, May 2008.

[24] Vinay Joseph and Gustavo de Veciana. Jointly Optimizing Multi-user Rate Adaptation for Video Transport over Wireless Systems: Mean-Fairness-Variability Tradeoffs. *Technical Report*, Jul. 2011. www.ece.utexas.edu/~gustavo/VariabilityAwareVideoRateAdapt.pdf.

[25] Vinay Joseph and Gustavo de Veciana. Variability aware network utility maximization. *arXiv:1111.3728v3 [cs.SY]*, 2011.

[26] Vinay Joseph, Gustavo de Veciana, and Ari Arapostathis. Resource Allocation: Realizing Mean-Variability-Fairness Tradeoffs. *Submitted to IEEE Transactions on Automatic Control*, 2013.

[27] F. P. Kelly, A. Maulloo, and D. Tan. Rate control for communication networks: Shadow prices, proportional fairness and stability. *Journal of Operation Research*, 1998.

[28] Taehyun Kim and M.H. Ammar. Optimal quality adaptation for scalable encoded video. *IEEE Journal on Selected Areas in Communications*, Dec. 2005.

[29] Harold J. Kushner and G. George Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[30] H.J. Kushner and P.A. Whiting. Convergence of proportional-fair sharing algorithms under general conditions. *IEEE Transactions on Wireless Communications*, Jul. 2004.

[31] Minghong Lin, Adam Wierman, Lachlan Andrew, and Eno Thereska. Online dynamic capacity provisioning in data centers. *Allerton Conference on Communication, Control and Computing*, 2011.

[32] Shihuan Liu, Lei Ying, and R. Srikant. Throughput-optimal opportunistic scheduling in the presence of flow-level dynamics. *IEEE/ACM Transactions on Networking*, 2011.

[33] Shie Mannor and John N. Tsitsiklis. Mean-Variance Optimization in Markov Decision Processes. *arXiv:1104.5601v1 [cs.LG]*, 2011.

[34] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, Oct. 2000.

[35] R.K.P. Mok, E.W.W. Chan, and R.K.C. Chang. Measuring the quality of experience of HTTP video streaming. *IFIP/IEEE International Symposium on Integrated Network Management*, 2011.

[36] A.K. Moorthy, K. Seshadrinathan, R. Soundararajan, and A.C. Bovik. Wireless video quality assessment: a study of subjective scores and objective algorithms. *IEEE Transactions on Circuits and Systems for Video Technology*, Apr. 2010.

[37] M.J. Neely, E. Modiano, and Chih-Ping Li. Fairness and optimal stochastic control for heterogeneous networks. *IEEE/ACM Transactions on Networking*, Apr. 2008.

[38] Leslie Real, James Ott, and Eva Silverfine. On the tradeoff between the mean and the variance in foraging: Effect of spatial distribution and color preference. *Ecology*, 1982.

[39] B. Sadiq and G. de Veciana. Throughput optimality of delay-driven maxweight scheduler for a wireless system with flow dynamics. *Allerton Conference on Communication, Control and Computing*, 2009.

[40] P. Seeling, M. Reisslein, and B. Kulapala. Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial. *IEEE Communications Surveys Tutorials*, 2004.

[41] S. Sen, J.L. Rexford, J.K. Dey, J.F. Kurose, and D.F. Towsley. Online smoothing of variable-bit-rate streaming video. *IEEE Transactions on Multimedia*, Mar. 2000.

[42] Kalpana Seshadrinathan and Alan Conrad Bovik. Automatic prediction of perceptual quality of multimedia signals–a survey. *International Journal of Multimedia Tools and Applications*, Jan. 2011.

[43] Srinivas Shakkottai and R. Srikant. Network optimization and control. *Foundations and Trends in Networking*, Jan. 2007.

[44] Georgi V. Smirnov. *Introduction to the Theory of Differential Inclusions*. American Mathematical Society, 2002.

[45] Matthew J. Sobel. Mean-variance tradeoffs in an undiscounted MDP. *Mathematics of Operations Research*, 1994.

[46] R. Soundararajan and A.C. Bovik. Video quality assessment by reduced reference spatio-temporal entropic differencing. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013.

[47] Roy A. Steiner and Michael F. Walter. The effect of allocation schedules on the performance of irrigation systems with different levels of spatial diversity and temporal variability. *Agricultural Water Management*, 1993.

[48] Alexander L. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, Aug. 2005.

[49] Alexander L. Stolyar. On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation. *Operations Research*, Jan. 2005.

[50] Alexander L. Stolyar. Greedy primal-dual algorithm for dynamic resource allocation in complex networks. *Queueing Systems*, Nov. 2006.

[51] K. Stuhlmuller, N. Farber, M. Link, and B. Girod. Analysis of video transmission over lossy channels. *IEEE Journal on Selected Areas in Communications*, Jun. 2000.

[52] Guan-Ming Su, Zhu Han, Min Wu, and K. J. Ray Liu. Joint uplink and downlink optimization for real-time multiuser video streaming over WLANs. *IEEE Transactions on Circuits and Systems for Video Technology*, Aug. 2007.

[53] Nikos Tsikriktsis and Janelle Heineke. The impact of process variation on customer dissatisfaction: Evidence from the U.S. domestic airline industry. *Decision Sciences*, 2004.

[54] P. van de Ven, S. Borst, and S. Shneer. Instability of maxweight scheduling algorithms. *IEEE INFOCOM*, 2009.

[55] Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers*, Nov. 2003.

[56] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, Apr. 2004.

[57] Zhou Wang, Ligang Lu, and Alan C. Bovik. Video quality assessment based on structural distortion measurement. *Signal Processing: Image Communication*, Feb. 2004.

[58] Ermin Wei, Asuman Ozdaglar, and Ali Jadbabaie. A distributed Newton method for Network Utility Maximization. *arXiv:1005.2633v2 [math.OC]*, 2011.

[59] Changhoon Yim and Alan C. Bovik. Evaluation of temporal variation of video quality in packet loss networks. *Signal Processing: Image Communication*, Jan. 2011.

[60] Xiaoqing Zhu and Bernd Girod. Analysis of multi-user congestion control for video streaming over wireless networks. *IEEE International Conference on Multimedia and Expo*, 2006.