Systems & Control Letters xxx (xxxx) xxx

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Utility maximization for asynchronous streaming of bufferable information flows

Vinay Joseph [a,*], Gustavo de Veciana [b]

[a] Department of Electronics and Communication Engineering, National Institute of Technology Calicut, India
[b] Department of Electrical and Computer Engineering, The University of Texas at Austin, United States of America

ARTICLE INFO

Article history:
Received 7 July 2022
Received in revised form 16 November 2022
Accepted 1 January 2023
Available online xxxx

Keywords:
Streaming
Information flow
Asynchronous
Rebuffering
Variability
Video
AR
VR

ABSTRACT

We consider optimizing the streaming of bufferable information flows to multiple clients sharing a network. The information flow to each client can be broken down into time segments where each segment is associated with a possibly varying quality/distortion-rate trade-off which can be adapted to network resources available and allocated to the client. The segments are downloaded, buffered and consumed sequentially by each client, and this proceeds in an *asynchronous* manner across the clients. Such settings are relevant to streaming of video and audio, and potentially to streaming of augmented reality and virtual reality content. We focus on jointly optimizing the network's resource allocation and clients' quality adaptation across segments so as to fairly optimize clients' Quality of Experience (QoE), while incorporating clients' sensitivity to rebuffering events caused when a client's buffer empties. We consider QoE models capturing trade-offs between clients' mean quality and temporal variability in quality. We present a *simple asymptotically optimal online* algorithm to solve the problem. It *distributes* the tasks of resource allocation to the network and quality adaptation to the respective clients. Further, it is *asynchronous* and is lightweight in terms of implementation overheads.

1. Introduction

We consider a setting where a network is shared by multiple clients. The network's resources (e.g., bandwidth) are shared by streams of *bufferable information flows* to the clients. A bufferable information flow to a client can be broken down into time *segments*. Each client downloads the segments sequentially into a buffer, and consumes downloaded segments at a constant rate whenever downloaded segments are available in the buffer. Downloaded segments that have not yet been consumed are stored at the client's buffer. Delivery of a segment to a client depends on the share of the network's finite resources allocated to the client. A segment can be delivered in one of multiple quality/distortion levels, and the amount of network resources required for the segment depends on the selected quality. More network resources are required to stream higher quality segments, and this 'quality-rate trade-off' can vary across segments for the same bufferable information flow.

An interesting aspect of this setting is *asynchrony*. That is, the network controller and clients operate 'at their own pace'. In particular, the completion instants of the segment downloads of clients do not need to align with those of other clients nor with the time epochs of network resource allocation. This asynchrony is well suited for modeling of asynchronous streaming of information flows. Such a setting matches, for example, the characteristics of video delivery to multiple devices over wireline/wireless networks, where the network might correspond to a base station shared by multiple end devices/clients consuming the content. In particular, the setting is compatible with Dynamic Adaptive Streaming over HTTP (DASH) based video streaming [1], wherein video is viewed as a sequence of short duration (e.g., seconds) video segments. Such settings may be relevant to streaming of augmented reality (AR) and virtual reality (VR) content too.

Our primary goal is to optimize the clients' Quality of Experience (QoE). To that end, we focus on solving the optimization problem OPT below:

$$\max \sum_{i \in \mathcal{N}} U_i\left(\text{Mean Quality}_i - \text{Quality Variability}_i\right)$$
subject to Rebuffering$_i$, and Network constraints,

where $\mathcal{N}$ is the set of clients, and U_i is a 'nice' function chosen to ensure desired fairness across clients' experienced QoE. The optimization variables for OPT are network resource allocation and clients' segment quality adaptation.

The objective function in OPT aims to capture QoE associated with the clients, where a client's QoE is expressed as being an increasing function of mean quality and decreasing function of

* Corresponding author.
E-mail address: vinayjoseph86@gmail.com (V. Joseph).

https://doi.org/10.1016/j.sysconle.2023.105455
0167-6911/

temporal variability in quality. While the role of mean quality (reduced distortion) in QoE is clear, the impact of temporal variability of quality is perhaps less and is often ignored in network resource allocation formulations. The detrimental impact of temporal variability on QoE of video streams can be significant (see [2–4]), and [2] 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.

Rebuffering constraints in the optimization problem, are related to rebuffering events which happen when the buffer of a client empties, and the client's consumption of information flow stalls. Rebuffering events have a significant impact on QoE. Indeed [5] 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 total time spent rebuffering.

Network constraints in OPT capture *time varying* constraints on network resource allocation, and can model wide range of real networks.

1.1. Main contributions

The main contribution of this paper is that it presents a general framework for optimization of streaming of bufferable information flows incorporating its inherent *asynchrony*. This asynchrony introduces several technical challenges, and they are discussed in Section 5.

In this paper, we present a *simple online* algorithm Resource Allocation and Quality Adaptation (RAQA), which solves OPT with strong optimality guarantees. In fact, our algorithm performs as well as omniscient optimal offline scheme which knows all about evolution of the network constraints and the quality-rate tradeoffs ahead of time. RAQA is also *asynchronous*, and using minimal communication, *distributes* the tasks of network resource allocation to network controller, and segment quality adaptation to respective clients.

1.2. Related work

The framework and algorithm developed in this paper enables developing of more holistic streaming solutions to important networking problems of today such as streaming of video (e.g., see [1], AR/VR (e.g., see [6]) and [7,8]) in advanced wireless networks.

There is a substantial body of work considering video streaming on which the present work draws, including in particular, [9–16] which utilize extensions of Network Utility Maximization (NUM) framework (see [17]). The main focus of [9,10] is real-time interactive video which present the challenge of meeting strict delivery deadlines. Papers [11,12] study video delivery optimization in wireless networks considering simpler QoE models, and do not explicitly incorporate rebuffering into their respective optimization frameworks, and instead control rebuffering through network congestion control. Using static QoE models, [14,15] study the resource allocation component of video delivery accounting for user dynamics. A major weakness of the aforementioned papers is the limited nature of the associated QoE models (that are essentially just the mean quality) and their lack of flexibility in managing/incorporating user preferences related to rebuffering.

While [13] presents a novel algorithm for realizing mean-variability tradeoffs for video streaming (see [18] for generalizations), the model involves a strong assumption of synchrony-the download of a segment of each video client starts at the beginning of a (network) slot and finishes by the end of the slot. This assumption on synchrony precludes any explicit control over rebuffering as it limits the ability of a video client to get ahead (by downloading more segments) during periods when channel is good and/or network is underloaded. Relaxed/different versions of this assumption can be found in the theoretical frameworks used in many previous papers (e.g., decision making in [11,12,16] is synchronous) as it facilitates an easier extension of tools from classical NUM framework. However, this assumption of synchrony is not ideal for DASH-based video clients in a wireless network that operate 'at their own pace'- downloading variable sized segments (with variable download times) one after the other. In this paper, we drop the assumption of synchrony which allows us to exploit opportunism across video clients' state of playback buffer (channels and features of video content like quality rate tradeoffs), and base our adaptation decision concerning a segment on network state information relevant to the download period of the segment. We also tackle the consequent novel technical challenges related to distributed asynchronous algorithms operating in a stochastic setting. Further, the rebuffering constraint in our asynchronous setting effectively induces a new type of constraint involving averages measured over two time scales.

From a theoretical perspective, our work relies heavily on an extension of results from the theory of asynchronous stochastic approximation presented in Chapter 12 of [19]. This extension is discussed in more detail in the proof sketch for Theorem 4. We also use extensions of several theoretical tools from [18,20] etc related to Network Utility Maximization (NUM). In summary, key novel elements of this work are the general QoE framework including incorporating sensitivity to temporal variability, rebuffering constraints involving two time-scales, asynchronous decision making and a generalization of results for asynchronous stochastic approximation in [19].

1.3. Organization of the paper

Section 2 introduces system model and assumptions. We formulate OPT in Section 1 as an *offline* optimization problem in Section 3.

In Section 4, we present an *online* algorithm RAQA which solves this optimization problem, and introduce its optimality properties. Proof of optimality of RAQA is discussed in Section 5. We conclude in Section 6.

2. System model

We shall first introduce some notation and conventions to be used throughout the paper. We use bold letters to denote vectors. Given a T-length sequence $(a(t))_{1\leq t\leq T}$ or a (infinite) sequence $(a(t))_{t\in\mathbb{N}}$, we let $(a)_{1:T}$ denote the T-length sequence $(a(t))_{1\leq t\leq T}$. For example consider a sequence $(\mathbf{a}(t))_{t\in\mathbb{N}}$ of vectors. Then $(\mathbf{a})_{1:T}$ denotes the T-length sequence containing the first T vectors of the sequence $(\mathbf{a}(t))_{t\in\mathbb{N}}$, and $(a_i)_{1:T}$ denotes the T-length sequence containing ith component of the first T vectors. Table 1 lists the key variables used in this paper. We develop our algorithmic framework by considering network supporting fixed set of bufferable information flows to clients $\mathcal{N}$ where $|\mathcal{N}| = N$. The network operates in a slotted manner with resources allocated for the duration of a slot τ_{slot} seconds. Slots are indexed by $k \in \{0, 1, 2\ldots\}$.

Allocation constraint: We assume that resource allocation is subject to time varying constraints. In each slot k, a network controller (e.g., a base station) allocates $\mathbf{r}_k = \left(r_{i,k}\right)_{i\in\mathcal{N}} \in \mathbb{R}_+^N$ bits (or $\mathbf{r}_k/\tau_{slot}$ bits per second) to the clients such that $c_k(\mathbf{r}_k) \leq 0$, where c_k is a real valued function modeling current constraints on network resource allocation. We refer to c_k as the *allocation constraint* in slot k. This function could be determined by various parameters like clients' SINR (Signal-to-Interference Noise Ratio).

Table 1
Summary of key notation used in this paper.

Variable	Description
$\mathcal{N}$	Set of clients served by network
N	Number clients served by network
i	Variable used to index clients in $\mathcal{N}$
$\overline{\beta}_i$	Client i's upper bound on fraction of time spent rebuffering
τ_{slot}	Slot duration in seconds
k	Variable used to index slots
c_k	Allocation constraint in slot k
$r_{i,k}$	Resource allocation to client i in slot k
l_i	Playback duration (in seconds) of each segment of client i
s, s_i	Variables used to index segments
$f_{i,s}$	QR trade-off associated with segment s of client i
$q_{i,s}$	Quality associated with segment s of client i
$m_{i,s}$	RAQA parameter for client i's mean quality till sth segment
$b_{i,k}$	RAQA parameter for client i's rebuffering risk in slot k

The available network resources may vary stochastically across slots. Hence we shall let C_k denote a randomly selected function corresponding to the allocation constraint in slot k, and thus c_k corresponds to a realization of such constraint for slot k.

We make the following assumptions on these allocation constraints:

Assumptions. C.1–C.3 (Time Varying Allocation Constraints)

C.1 $(C_k)_{k\in\mathbb{N}}$ 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 $[0, r_{\max}]^N$ with $c(\mathbf{0}) \leq 0$ and

$$\min_{\mathbf{r}\in[0,r_{\max}]^N} c(\mathbf{r}) < 0. \tag{1}$$

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 $c(\mathbf{r}) \leq 0$, we have $r_i \leq r_{\max}$ for each $i \in \mathcal{N}$.

We denote the marginal distribution of this process by $\left(\pi^{\mathcal{C}}(c)\right)_{c\in\mathcal{C}}$. Without loss of generality, we assume that $\pi^{\mathcal{C}}(c) > 0$ for each $c \in \mathcal{C}$. Note that we are restricting ourselves to settings with convex capacity regions $\left\{\left(r_{i,k}\right)_{i\in\mathcal{N}} \in \mathbb{R}_+^N : c_k(\mathbf{r}_k) \leq 0\right\}$ due to the convexity assumption in C.2. This model captures a fairly general class of allocation constraints, including, for example, time-varying capacity constraints associated with bandwidth allocation at a shared wireless base station. Further, we require that the resources allocated to each client $i \in \mathcal{N}$ in each slot should be at least $r_{i,\min}$ where $r_{i,\min} > 0$ (can be relaxed to each client being guaranteed a strictly positive amount of resource allocation over a fixed (large) number of slots).

Information flows' Quality Rate (QR) tradeoffs: We assume information flows are modeled as sequence of segments which are streamed to, buffered and consumed by the associated client. Segments are indexed using variables like s, s_i etc taking values in $\{0, 1, 2, \ldots\}$. We let l_i denote the playback duration, in seconds, of each segment of client i (extensions to segment-index-dependent segment durations are in [21]). Each segment of the flow may have multiple representations realizing possible Quality-Rate (QR) tradeoffs. Specifically the QR tradeoff for the sth segment of client i is captured by a convex function $f_{i,s}$ – to obtain a representation of quality $q_{i,s}$, the client will need to download a file of size $l_i f_{i,s}(q_{i,s})$. Convexity is a reasonable assumption for video [1], where increases in the rate typically lead to diminishing marginal gains in the video quality/distortion.

As with the resource allocation constraints, we shall consider a setting where the QR tradeoff associated with the sth segment of client i is modeled by a random function $F_{i,s}$ satisfying the following assumptions:

Assumptions. QR.1–QR.2 on QR Tradeoffs

QR.1 For each client $i \in \mathcal{N}$ we assume $\left(F_{i,s}\right)_{s\geq 0}$ is a *stationary ergodic* process taking values in a set $\mathcal{F}_i$.

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

As indicated in Assumption QR.1, we model the evolution of QR tradeoffs of each client $i \in \mathcal{N}$ as a stationary ergodic process. Let $\left(\pi^{\mathcal{F}_i}(f_i)\right)_{f_i\in\mathcal{F}_i}$ denote the associated marginal distribution. Without loss of generality, we assume that $\pi^{\mathcal{F}_i}(f_i) > 0$ for each $f_i \in \mathcal{F}_i$. Let $f_{\min} := \min_{\{i\in\mathcal{N}, f_i\in\mathcal{F}_i\}} f_i(0)$ which is strictly positive from QR.2, and this gives a lower bound on segment compression rates. Even at zero quality, there will be 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.

QoE model: A client's Quality of Experience (QoE), depends on the quality of the segment representations, $(q_i)_{1:S}$, downloaded by a client i on the condition that a rebuffering related constraint (discussed next) is met. QoE models are typically complex and context dependent. We shall adopt a simple[1] model motivated by the discussion in Section 1 and the model in [2]. Let $m_i^S((q_i)_{1:S})$ and $\mathrm{Var}_i^S((q_i)_{1:S})$ denote mean quality and temporal variance in quality respectively associated with the first S segments downloaded by the client i, i.e.,

$$m_i^S\left((q_i)_{1:S}\right) := \frac{\sum_{s=1}^S q_{i,s}}{S},$$

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

We model the QoE of client i for these S segments as

$$e_i^S\left((q_i)_{1:S}\right) = m_i^S\left((q_i)_{1:S}\right) - \eta_i \mathrm{Var}^S\left((q_i)_{1:S}\right), \tag{2}$$

where $\eta_i > 0$ scales the penalty for temporal variability in quality.

Our objective function capturing clients' QoE is

$$\phi_S\left((\mathbf{q})_{1:S}\right) := \sum_{i\in\mathcal{N}} e_i^S\left((q_i)_{1:S}\right). \tag{3}$$

Here, we have set $U_i(.)$ appearing in OPT in Section 1 to $U_i(e) = e$. In [21], we discuss extensions to concave $U_i(.)$ which provide more flexibility in imposing QoE fairness across users, and consider more general variability penalties involving non-linear functions of $\mathrm{Var}^S\left((q_i)_{1:S}\right)$.

Rebuffering constraints: Let $\kappa > 0$ and let $K_S = \lceil \kappa S \rceil$. In the sequel, we shall focus on stationary resource allocation policies such that for each client i, $\frac{1}{K_S}\sum_{k=1}^{K_S} r_{i,k}$ converges.

Then, a good estimate (for large S) for the time required by client i to download the first S segments is

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

which is the ratio of the cumulative size of S segments (i.e., $l_i \sum_{s=1}^S f_{i,s}(q_{i,s})$) to the per slot resource allocation estimate (i.e., $\frac{1}{\tau_{slot} K_S}\sum_{k=1}^{K_S} r_{i,k}$). Note that since we consider settings where

[1] See [21] for extensions to more general QoE models.

the per slot resource allocation converges, it follows that $T_{i,S}(.)$ is an asymptotically (as S goes to infinity) accurate estimate of the time required by client i to download the first S segments. Thus, the following is an asymptotically accurate estimate for the percentage of time that client i is rebuffering while watching the S segments:

$$\beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right) := \frac{T_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right) - \sum_{s=1}^{S} l_i}{\sum_{s=1}^{S} l_i}. \tag{4}$$

The denominator in (4) corresponds to the total playback duration $\sum_{s=1}^{S} l_i$ associated with the S segments. The numerator term is an estimate of time spent rebuffering, as it is the difference of the estimate for time required for download of the first S segments and the total playback duration associated with the S segments. Note that we allow $\beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right)$ can also take negative values which happens when segments are being downloaded at rate higher than the rate at which they are consumed. We express the rebuffering constraint as

$$\beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right) \le \overline{\beta}_i, \ \forall\, i \in \mathcal{N}, \tag{5}$$

where each client i specifies an upper bound $\overline{\beta}_i > -1$ on the fraction 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.

3. Offline optimization formulation

Given the system model in the previous section, one can now formalize the *offline* optimization problem OPT(S) corresponding to maximizing the clients' sum QoE. The maximization is over joint selection of the clients' quality adaptation (i.e., finding $\left((q_i)_{1:S}\right)_{i\in\mathcal{N}}$) and the network's resource allocation (i.e., finding $(\mathbf{r})_{1:K_S}$). In the offline setting, we assume that $(c_k)_k$ and $\left(f_{i,s}\right)_s$ for each client $i \in \mathcal{N}$ are known ahead of time. The offline optimization OPT(S) is given below.

OPT(S)

$$\max_{(\mathbf{q})_{1:S}, (\mathbf{r})_{1:K_S}} \quad \phi_S\left((\mathbf{q})_{1:S}\right)$$

$$\begin{aligned} \text{subject to} \quad & 0 \le q_{i,s} \le q_{\max}\ \forall\, s \in \{1,\ldots,S\}, \forall\, i \in \mathcal{N}, \\ & r_{i,k} \ge r_{i,\min},\ \ \forall\, k \in \{1,\ldots,K_S\}, \forall\, i \in \mathcal{N}, \\ & c_k(\mathbf{r}_k) \le 0, \ \ \forall\, k \in \{1,\ldots,K_S\}, \\ & \beta_{i,S}\left((q_i)_{1:S}, (r_i)_{1:K_S}\right) \le \overline{\beta}_i, \forall\, i \in \mathcal{N}. \end{aligned} \tag{6}$$

We will further require following assumption to ensure strict feasibility, and this is used in later sections.

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

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

We assume that the optimization problem OPT(S) is feasible (sufficient conditions are discussed in [21]). Let ϕ_S^{opt} denote the optimal value of objective function of OPT(S).

In practice, solving OPT(S) directly is impossible (except for trivial cases) since we need to know $(c_k)_k$ and $\left(f_{i,s}\right)_s$ ahead of time. Further, it is also computationally prohibitive as the optimization would be over $O(NS)$ variables. Thus, the challenge is to overcome these two hurdles, i.e., to find a *simple* and *online* algorithm that is near optimal.

4. Online algorithm for resource allocation and quality adaptation

In this section, we propose an algorithm for joint Resource Allocation and Quality Allocation denoted RAQA for short. It involves solving network-level and client-level optimization problems RA($\mathbf{b}$, c) and QA$_i$($\boldsymbol{\theta}_i, f_i$) which are discussed next.

Resource allocation problem RA($\mathbf{b}$, c): For $\mathbf{b} \in \mathbb{R}^N$ and allocation constraint $c \in \mathcal{C}$, the (convex) optimization problem RA($\mathbf{b}$, c) for resource allocation is

$$\max_{\mathbf{r}} \left\{ \sum_{i\in\mathcal{N}} h_i^B(b_i)\, r_i :\ c(\mathbf{r}) \le 0,\ r_i \ge r_{i,\min}\ \forall i \in \mathcal{N} \right\}, \tag{7}$$

where $h_i^B(.)$ is a non-negative valued Lipschitz continuous function such that $\lim_{b\to\infty} h_i^B(b) = \infty$, $h_i^B(b_i) = 0$ for all $b_i \le \underline{b}$ for some constant $\underline{b}$ (typically set as zero or small negative numbers), and is strictly increasing for $b_i \ge \underline{b}$. Simple examples of functions satisfying these conditions are max(b, 0), max(b^2, 0) etc. Let $\mathcal{R}^*(\mathbf{b}, c)$ denote the set of optimal solutions to RA($\mathbf{b}$, c). Also, let $\phi^R(\mathbf{r}, \mathbf{b}, c)$ denote the objective function of RA($\mathbf{b}$, c).

Client quality adaptation problem QA$_i$($\boldsymbol{\theta}_i, f_i$): For $m_i \in [0, q_{\max}]$, $b_i \in \mathbb{R}$ and $\boldsymbol{\theta}_i = (m_i, b_i)$, let

$$\phi^Q(q_i, \boldsymbol{\theta}_i, f_i) = q_i - \eta_i (q_i - m_i)^2 - \frac{h_i^B(b_i)}{\left(1 + \overline{\beta}_i\right)} f_i(q_i) \tag{8}$$

for QR tradeoff f_i. The optimization problem QA$_i$($\boldsymbol{\theta}_i, f_i$) associated with quality adaptation of client i is given below:

$$\max_{q_i} \left\{ \phi^Q(q_i, \boldsymbol{\theta}_i, f_i) :\ 0 \le q_i \le q_{\max} \right\}.$$

The optimization problem QA$_i$($\boldsymbol{\theta}_i, f_i$) is convex with strictly concave objective function, and thus has a unique solution denoted as $q_i^*(\boldsymbol{\theta}_i, f_i)$.

RAQA algorithm: We will further elaborate on the intuition underlying the two types of optimization problems introduced above after presenting RAQA framework.

Let s_i be an indexing variable keeping track of the segment that client i is currently downloading. Let $\epsilon > 0$,

$$\mathcal{H}^{(i)} = \left\{ (m_i, b_i) \in \mathbb{R}^3 : 0 \le m_i \le q_{\max},\ b_i \ge \underline{b} \right\},$$

and let $[x]_y = \max(x, y)$ for $x, y \in \mathbb{R}$.

RAQA

Initialization: Let $\left(m_{i,0}, b_{i,0}\right) \in \mathcal{H}^{(i)}$ for each $i \in \mathcal{N}$.

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

ALLOCATE: At the beginning of slot k, network controller allocates resources $\mathbf{r}_k^*$ choosing any solution to RA($\mathbf{b}_k$, c_k). Update $\mathbf{b}_k$ as follows:

$$b_{i,k+1} = b_{i,k} + \epsilon \left(\frac{\tau_{slot}}{\left(1 + \overline{\beta}_i\right)} \right). \tag{9}$$

ADAPT: In slot k, if any client $i \in \mathcal{N}$ finishes download of s_i th segment, download of segment $s_i + 1$ is started immediately. Let $\boldsymbol{\theta}_{i,s_i} = (m_{i,s_i}, b_{i,k+1})$. For segment $s_i + 1$ of client i, the client selects representation with quality $q_i^*(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$ (i.e., optimal solution to QA$_i$($\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1}$)), denoted as q_{i,s_i+1}^* for brevity. Update parameters m_{i,s_i+1}, $b_{i,k+1}$ and s_i as follows:

$$m_{i,s_i+1} = m_{i,s_i} + \epsilon\left(q_{i,s_i+1}^* - m_{i,s_i}\right), \tag{10}$$

$$b_{i,k+1} = \left[b_{i,k+1} - \epsilon\,(l_i) \right]_{\underline{b}}, \tag{11}$$

$$s_i = s_i + 1.$$

Thus, RAQA's resource allocation is done at the beginning of each slot k, allocating $r^*_{i,k}$ to client i. RAQA's quality adaptation for segment s_i+1 of client i is carried out immediately after completion of download of segment s_i, and involves selecting representation with quality $q^*_i(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$. Thus, the quality adaptation is asynchronous, i.e., adaptation related decisions about a segment are made by a client only at the completion of download of previous segment.

The update Eq. (10) associated with the parameter m_{i,s_i} is similar to update rules used for tracking EWMA (Exponentially Weighted Moving Averages), and ensures that m_{i,s_i} tracks the mean quality of client i. Consider the evolution of $b_{i,k}$ which is updated in both (9) and (11) ignoring the operator $[.]_{\underline{b}}$ and with initialization to zero. Hence, at some time t seconds (or $k = t/\tau_{slot}$ slots) after starting the video,

$$\frac{b_{i,k} - b_{i,0}}{\epsilon} \approx \frac{t}{\left(1+\overline{\beta}_i\right)} - L^D_i(t),$$

where $L^D_i(t)$ is the playback duration of segments downloaded up to time t. This sheds light on the role of $b_{i,k}$ as an indicator of risk of violation of rebuffering constraint in (6) for client i. For $\overline{\beta}_i = 0$ and small enough $\underline{b}$, $(b_{i,k} - b_{i,0})/\epsilon$ is equal to $(t - L^D_i(t))$ which is equal to *negative* of the duration of content in playback buffer (if there is any).

The above discussion about the roles of m_{i,s_i} and $b_{i,k}$ now make the formulations of optimization problems RA($\mathbf{b}, c$) and QA$_i(\boldsymbol{\theta}_i, f_i)$ intuitively clearer. The objective function (8) includes a term $\left(q_i - m_{i,s}\right)^2$ ensuring that an optimal solution to QA$_i(\boldsymbol{\theta}_i, f_i)$ is not too far from $m_{i,s}$. Since $m_{i,s}$ tracks current estimate of mean quality, this avoids high variance in quality. Next, note that the term $\frac{h^B_i(b_i)}{(1+\overline{\beta}_i)} f_i\left(q_i\right)$ in (8) penalizes quality choices leading to large segment sizes when $b_{i,k+1}$ is high (i.e., there is higher risk of violation of rebuffering constraints). Further, relatively large $b_{i,k}$ results in higher allocation to client i (see (7)). Also note that we can control the sensitivity of RAQA to a higher risk of rebuffering by appropriately choosing $\left(h^B_i(.)\right)_{i\in\mathcal{N}}$.

For each $i \in \mathcal{N}$, parameters $\left(m_{i,s_i}, b_{i,k}\right)$ are learnt/updated by client i. The network controller only needs to know $\mathbf{b}_k$ for carrying out resource allocation in slot k. and this can be achieved using minimal signaling (see Section 4.2). Further, strong optimality of RAQA (see Theorem 1) points to the critical role $\left(b_{i,k}\right)_{i\in\mathcal{N}}$ in carrying optimal joint quality adaptation and resource allocation using minimal signaling. In particular, $\left(b_{i,k}\right)_{i\in\mathcal{N}}$ carries almost all the information about the 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 client i needs to know about the resource allocation (to the client).

4.1. Optimality of RAQA

The following theorem is the main optimality result for RAQA, and we discuss key steps of our proof in Section 5.

Theorem 1. *Under the assumptions in the previous sections,*

(a) Feasibility: *RAQA asymptotically satisfies the rebuffering constraints, i.e., for each $i \in \mathcal{N}$*

$$\limsup_{S\to\infty} \beta_{i,S}\left(\left(q^*_i\right)_{1:S}, \left(r^*_i\right)_{1:K_S}\right) \le \overline{\beta}_i. \tag{12}$$

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

$$\lim_{S\to\infty}\lim_{\epsilon\to 0}\left(\phi_{S_\epsilon}\left(\left(\mathbf{q}^*\right)_{1:S_\epsilon}\right) - \phi^{opt}_{S_\epsilon}\right)$$

converges to zero in probability.

Here C_k and $F_{i,s}$ are random functions corresponding to c_k and $f_{i,s}$ respectively. Recall that, under RAQA, q^*_{i,s_i} is the quality associated with segment s_i of client i (and the notation used in this result is described at the beginning of Section 2). This result tells us that the difference in performance (according to definition (3)) of the *online* algorithm RAQA (i.e., $\phi_{S_\epsilon}\left(\left(\mathbf{q}^*\right)_{1:S_\epsilon}\right)$) and that of the optimal *offline* scheme goes to zero for long enough videos and small enough ϵ. Recall that $\phi^{opt}_{S_\epsilon}$ is the optimal value of 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 $\left(f_{i,s}\right)_s$ ahead of time.

4.2. Key features and implementation of RAQA

Next, we summarize the key features of RAQA.

Optimality: RAQA carries out 'cross-layer' joint optimization of resource allocation and quality adaptation, with strong optimality guarantees given in Theorem 1.

Online: RAQA is an online algorithm as it only uses *current* information, i.e., network controller only needs to know the allocation constraint c_k to carry out resource allocation for slot k, and client i only requires the QR tradeoff $f_{i,s}$ for quality adaptation of segment s.

Simple: QA$_i(\boldsymbol{\theta}_i, f_i)$ is a scalar convex optimization problem. RA($\mathbf{b}, c$) is an N variable convex optimization problem.

Distributed implementation: RAQA can be implemented in a *distributed* manner with minimal signaling since quality adaptation is *client driven* and for the resource allocation, the network controller needs to only know $\mathbf{b}_k$. Each client can send a signal to the network controller indicating the end of each segment download, which can trigger (11) at the network controller. The network controller could obtain information about allocation constraints through Channel Quality Information (CQI) feedback from the network, and clients could obtain their respective QR tradeoffs using application layer information exchange. Further, the asynchronous nature of RAQA ensures that the clients can work at their own pace. Thus, the adaptation prescribed in RAQA is entirely *client driven* requiring no assistance from the network controller.

5. Proof of optimality of RAQA

This section develops the proof of Theorem 1, which is the main optimality result for RAQA. Due to space constraints, we have included shortened proofs for some of the intermediate results (focusing on key ideas involved) and detailed proofs can be found in [21].

Outline of the proof: To provide a big picture for overall proof, we begin with an outline below highlighting the key steps:

(1) We begin by discussing about some key properties of RAQA (see Lemmas 1–3).

(2) *RAQA is optimal if the underlying estimated parameters were to converge to optimal parameter set:* In particular, in Section 5.1, we study an auxiliary optimization problem OPT-STAT and obtain Theorem 2 which suggests that our key Theorem 1 will follow if RAQA estimated parameters converge to an optimal set.

(3) *RAQA's auxiliary differential inclusion converges to the optimal parameter set:* In Section 5.2, we study an auxiliary differential inclusion ((27)–(31)) which evolves according to average dynamics of RAQA, and show in Theorem 3 that the differential inclusion converges to the optimal parameter set.

(4) *RAQA parameters also converge to the optimal parameter set:* In Section 5.3, we view RAQA's update equations ((10)–(11) and (23)– (25)) as an asynchronous stochastic approximation update (see, e.g., [19] for reference), and relate this stochastic approximation update to the auxiliary differential inclusion (in (27)–(31)). We then use this relationship and convergence of the auxiliary differential inclusion (from Section 5.2) to show that RAQA's parameters also converge to the optimal parameter set.
(5) Finally, a proof of Theorem 1 is given in Sub Section 5.4.

We shall thus begin by discussing some key properties of RAQA. The optimization problem $\mathrm{RA}(\mathbf{b}, c)$ is convex, and using Assumption-SF, we can show that it satisfies Slater's condition (see e.g., [22]). Thus, KKT conditions are necessary and sufficient for optimality. The optimization problem $\mathrm{QA}_i(\boldsymbol{\theta}_i, f_i)$ is also convex and satisfies Slater's condition (since the constraints are all linear), and thus, KKT conditions are necessary and sufficient for optimality.

The next result states that the parameters in RAQA stay in a compact set and in particular, points out that the parameters $b_{i,k}$ can be uniformly bounded.

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

Proof. The boundedness of $m_{i,.}$ and lower bound for $b_{i,.}$ can be proved using boundedness of the quantities involved in (10) and (11) respectively. Proving an upper bound for $b_{i,.}$ involves noting that larger $b_{i,.}$ values force the selection of lower quality/size segments, and $b_{i,.}$ stops increasing once the segment size is small enough (based on the role of $b_{i,.}$ in $\mathrm{QA}_i(\boldsymbol{\theta}_i, f_i)$). ■

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

Lemma 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. Parts (a), (b) and (c) follow respectively from Theorem 2.2, Theorem 2.4 and Theorem 2.1 in [23]. ■

In the next result, we discuss concavity and differentiability properties of the optimal value of $\mathrm{QA}_i(\boldsymbol{\theta}_i, f_i)$.

Lemma 3. *The following hold for each $i \in \mathcal{N}$ and $f_i \in \mathcal{F}_i$.*

(a) The optimal value of $\mathrm{QA}_i(\boldsymbol{\theta}_i, f_i)$, i.e., $\phi^Q\left(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i\right)$, is a strictly concave function of m_i (with b_i and d_i fixed).

(b) The partial derivative of $\phi^Q\left(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i\right)$ with respect of m_i is given by:

$$\frac{\partial \phi^Q\left(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i\right)}{\partial m_i} = 2\eta_i\left(q_i^*(\boldsymbol{\theta}_i, f_i) - m_i\right). \tag{13}$$

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

$$\phi^Q\left(q_i^*\left(\boldsymbol{\theta}_i^{(m)}, f_i\right), \boldsymbol{\theta}_i^{(m)}, f_i\right) < \phi^Q\left(q_i^*(\boldsymbol{\theta}_i, f_i), \boldsymbol{\theta}_i, f_i\right) + 2\eta_i(m - m_i)\left(q_i^*(\boldsymbol{\theta}_i, f_i) - m_i\right).$$

Proof. Part (a) follows from Proposition 2.8 from [24]. Part (b) follows from Theorem 4.1 in [25]. Part (c) follows from strict concavity in (a) and (b). ■

5.1. RAQA is optimal if the underlying estimated parameters were to converge to optimal parameter set

Next we develop an optimal solution which corresponds to a *stationary* policy, i.e., a policy for which the allocation and quality adaptation decisions depend solely on the *current* state determined by the current allocation constraint or QR tradeoffs respectively. At first sight, it may not be apparent how a stationary policy utilizing such limited information (i.e., just the state) could help towards solving $\mathrm{OPT}(S)$ (which involves complex terms like temporal variance, and (6)). Intuitively this is due to the fact that the objective and constraint functions in $\mathrm{OPT}(S)$ depend on time and segment averages of the quantities of interest, e.g. mean and variance of segment quality over time. Thus, in the stationary ergodic regime, the averages involved in $\mathrm{OPT}(S)$ can be rewritten in terms of ensemble averages. Additionally, note that we will consider stationary policies where resource allocation is not directly dependent on QR-tradeoffs and quality adaptation is not directly dependent on current allocation constraint. Still we will see that such a policy performs well. This is explored in Lemma 4(b) which suggests that all the information required to jointly coordinate resource allocation and quality adaptation can be encapsulated in few shared parameters.

Recall (see Section 2) that $(C_k)_k$ is a stationary ergodic random process with marginal distribution $\left(\pi^{\mathcal{C}}(c)\right)_{c\in\mathcal{C}}$. We let C^π denote a random function with distribution $\left(\pi^{\mathcal{C}}(c) : c \in \mathcal{C}\right)$. Also, recall that for each $i \in \mathcal{N}$, $\left(F_{i,s}\right)_{s\geq 0}$ is a stationary ergodic process with marginal distribution $\left(\pi^{\mathcal{F}_i}(f_i)\right)_{f_i\in\mathcal{F}_i}$. We let F_i^π denote a random function with distribution $\left(\pi^{\mathcal{F}_i}(f_i)\right)_{f_i\in\mathcal{F}_i}$.

Consider a stationary policy with $(\mathbf{r}(c))_{c\in\mathcal{C}}$ being a vector (of vectors) representing the allocation of resources $\mathbf{r}(c) \in \mathbb{R}^N$ for each possible instance of resource constraints, i.e., $c \in \mathcal{C}$. The above represents a slight abuse of notation as earlier we let $\mathbf{r}(t)$ denote the allocation to the clients in slot t. One can easily differentiate between these based on the context in which they are being discussed. Also, for such a stationary policy, let $q_i(f)$ denote the quality associated with a segment of client i when $f \in \mathcal{F}_i$. Mimicking the definition of $\phi_S\left((\mathbf{q})_{1:S}\right)$, $m_i^S\left((q_i)_{1:S}\right)$ and $\mathrm{Var}_i^S\left((q_i)_{1:S}\right)$ in Section 3, we let

$$\phi_\pi\left(\left((q_i(f_i))_{f_i\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}\right) = \sum_{i\in\mathcal{N}}\left(\mathrm{Mean}\left(q_i\left(F_i^\pi\right)\right) - \eta_i \mathrm{Var}\left(q_i\left(F_i^\pi\right)\right)\right), \tag{14}$$

$$\mathrm{Mean}\left(q_i\left(F_i^\pi\right)\right) = \mathbb{E}\left[q_i\left(F_i^\pi\right)\right],$$

$$\mathrm{Var}\left(q_i\left(F_i^\pi\right)\right) = \mathbb{E}\left[\left(q_i\left(F_i^\pi\right) - \mathrm{Mean}\left(q_i\left(F_i^\pi\right)\right)\right)^2\right].$$

Let us now consider the optimization problem OPTSTAT.

OPTSTAT

$$\max_{\left((q_i(f_i))_{f_i\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, (\mathbf{r}(c))_{c\in\mathcal{C}}} \quad \phi_\pi\left(\left((q_i(f_i))_{f_i\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}\right) \tag{15}$$

$$\text{subject to } c(\mathbf{r}(c)) \leq 0, \ \forall\, c \in \mathcal{C}, \tag{16}$$

$$0 \leq q_i(f_i) \leq q_{\max}, \ \forall\, f_i \in \mathcal{F}_i, \ \forall\, i \in \mathcal{N},$$

$$r_i(c) \geq r_{i,\min}, \ \forall\, c \in \mathcal{C}, \ \forall\, i \in \mathcal{N},$$

$$\frac{\mathbb{E}\left[F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right]}{\left(1 + \overline{\beta}_i\right)} \leq \frac{\mathbb{E}\left[r_i\left(C^\pi\right)\right]}{\tau_{slot}}, \ \forall\, i \in \mathcal{N}. \tag{17}$$

This optimization problem is obtained by replacing the time and segment averages of various quantities in OPT(S) with the expected value of the corresponding quantities evaluated under the stationary distribution of $(C_k)_k$ and $\left(F_{i,s}\right)_{s\geq 0}$ for each $i \in \mathcal{N}$. Note that in the constraint $c\left(\mathbf{r}\left(c\right)\right) \leq 0$ given in (16), 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. Similarly, in the term $F_i^\pi\left(q_i\left(F_i^\pi\right)\right)$, the argument F_i^π serves as an index whereas $F_i^\pi\left(.\right)$ is the (random) function.

We can show that OPTSTAT is a convex optimization problem satisfying Slater's condition. Further, we can show that the optimal quality choices obtained by solving OPTSTAT are unique and we denote them by $\left(\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}$. Let $\left(\left(\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^\pi\left(c\right)\right)_{c\in\mathcal{C}}\right)$ be an optimal solution to OPTSTAT, and let $\mathbf{b}^\pi$ denote the associated Lagrange multipliers for the constraints (17). Since OPTSTAT is a convex optimization problem satisfying Slater's condition, we can conclude that the KKT conditions are necessary and sufficient for optimality. For each $i \in \mathcal{N}$, let

$$m_i^\pi = \mathbb{E}\left[q_i^\pi\left(F_i^\pi\right)\right], \tag{18}$$

$$v_i^\pi = \mathrm{Var}\left(q_i^\pi\left(F_i^\pi\right)\right), \tag{19}$$

$$\sigma_i^\pi = \mathbb{E}\left[F_i^\pi\left(q_i^\pi\left(F_i^\pi\right)\right)\right]. \tag{20}$$

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

$$\begin{aligned}\mathcal{X}^\pi = \{(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi) : &\text{ there is an optimal solution} \\ &\left(\left(\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^\pi\left(c\right)\right)_{c\in\mathcal{C}}\right) \text{ to OPTSTAT with} \\ &\rho_i^\pi = \frac{\mathbb{E}\left[r_i^\pi\left(C^\pi\right)\right]}{\tau_{slot}} \text{ for each } i \in \mathcal{N}, \text{ and with} \\ &\mathbf{b}^\pi \text{ as the associated optimal Lagrange multipliers} \\ &\text{for constraints (17) respectively}\}.\end{aligned} \tag{21}$$

In the next result, we present three useful properties of any optimal solution to OPTSTAT. Part (a) below provides a client level optimality result which essentially suggests that we can decouple the quality adaptation of the 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 client $i \in \mathcal{N}$ is itself an optimal solution to an optimization problem which can be solved by the client i. Part (b) points out that we only need to know a few parameters (specifically, $\mathbf{b}^\pi$) 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 when the RAQA parameters $\boldsymbol{\theta}_{i,s}$ of client i are in the set $\mathcal{H}_i^*$ defined as

$$\begin{aligned}\mathcal{H}_i^* := \Big\{\left(m_i^\pi, \left(h_i^B\right)^{-1}\left(b_i^\pi\right)\right) : &\text{ there exists some } \boldsymbol{\rho} \in \mathbb{R}^N \\ &\text{such that } (\boldsymbol{\rho}, \mathbf{b}^\pi) \in \mathcal{X}^\pi\Big\},\end{aligned} \tag{22}$$

the RAQA can provide optimal quality choices for OPTSTAT.

Lemma 4. *For parts (a) and (b) of this result, suppose* $(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi) \in \mathcal{X}^\pi$ *and let the associated optimal solution be* $\left(\left(\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}\right)_{i\in\mathcal{N}}, \left(\mathbf{r}^\pi\left(c\right)\right)_{c\in\mathcal{C}}\right)$.

(a) For each $i \in \mathcal{N}$, $\left(q_i^\pi\left(f\right)\right)_{f\in\mathcal{F}_i}$ *is the unique optimal solution to the following optimization problem*

$$\max_{\left(\left(q_i(f)\right)_{f\in\mathcal{F}_i}\right)} \mathbb{E}\left[q_i\left(F_i^\pi\right)\right] - \eta_i Var\left(q_i\left(F_i^\pi\right)\right)$$
$$- \sum_{i\in\mathcal{N}} \frac{b_i^\pi}{\left(1+\overline{\beta}_i\right)} \mathbb{E}\left[F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right],$$
$$s.t.\ 0 \leq q_i(f) \leq q_{\max},\ \forall f \in \mathcal{F}_i.$$

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

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

(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. Parts (a) and (b) can be shown using KKT conditions for OPTSTAT, and applying them to the optimization problems in parts (a) and (b). Part (c) follows from strict concavity of the objective function in part (a). ■

The next result states that the performance of RAQA (measured in terms of $\phi_S(.)$ defined in (3)) with its parameters $\boldsymbol{\theta}_{i,s}$ picked from the set $\mathcal{H}_i^*$ is optimal.

Theorem 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. Proof details are omitted for brevity (see Theorem 3.1 [21] for details). The proof involves constructing a feasible solution to OPTSTAT (i.e., a stationary policy) using optimal temporal solution, and comparing its performance against that of the optimal stationary policy, and arguing that both lead to same performance by utilizing optimality of respective formulations, and stationarity and ergodicity of underlying processes $\left(F_{i,s}, L_{i,s}\right)_{s\geq 0}$. ■

5.2. RAQA's auxiliary differential inclusion converges to the optimal parameter set

Theorem 2 suggests that we can prove Theorem 1 if we can show that the updates (10)–(11) of RAQA guide the parameters $\left(\boldsymbol{\theta}_{i,s}\right)_{s\geq 1}$ of client i to $\mathcal{H}_i^*$ for each client $i \in \mathcal{N}$. Hence, we next show that RAQA's 'learning component' (i.e., updates (9)–(11)) guide its parameters to the optimal set (i.e., $\prod_{i\in\mathcal{N}} \mathcal{H}_i^*$). To this end, we study an *auxiliary differential inclusion* which evolves according to average dynamics of RAQA, and study its convergence.

For the rest of this section, we also consider the evolution of auxiliary parameters $\left(v_{i,s_i}\right)_{s_i\geq 1}$, $\left(\sigma_{i,s_i}\right)_{s_i\geq 1}$ and $\left(\rho_{i,k}\right)_{k\geq 1}$ associated with RAQA. We update v_{i,s_i} and σ_{i,s_i} based on the quality q_{i,s_i+1}^* (shorthand for $q_i^*(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$ where $\boldsymbol{\theta}_{i,s_i} = (m_{i,s_i}, b_{Q,i,s_i})$) chosen by RAQA for (s_i+1)th segment of client $i \in \mathcal{N}$ as follows:

$$v_{i,s_i+1} = v_{i,s_i} + \epsilon\left(\left(q_{i,s_i+1}^* - m_{i,s_i}\right)^2 - v_{i,s_i}\right), \tag{23}$$

$$\sigma_{i,s_i+1} = \sigma_{i,s_i} + \epsilon\left(f_{i,s_i}\left(q_{i,s_i+1}^*\right) - \sigma_{i,s_i}\right). \tag{24}$$

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

$$\rho_{i,k+1} = \rho_{i,k} + \epsilon\left(\frac{r_{i,k}^*}{\tau_{slot}} - \rho_{i,k}\right)\ \forall i \in \mathcal{N}. \tag{25}$$

Thus, $\boldsymbol{\rho}_k$ tracks the mean resource allocation per unit time to clients.

Thus, the auxiliary parameters do not affect the allocation or adaptation in RAQA. The evolution of parameters $(\mathbf{m}_s, \mathbf{v}_s,$

$\left(\mathbf{b}_k, \boldsymbol{\sigma}_s, \boldsymbol{\rho}_k\right)_{s,k}$ does depend on RAQA's resource allocation ($\mathbf{r}_k^*$) and quality adaptation ($q_i^*(\boldsymbol{\theta}_{i,s_i}, f_{i,s_i+1})$).

Using Lemma 1 and assumptions in Section 2, we can show that the $\left(\mathbf{m}_s, \mathbf{v}_s, \mathbf{b}_k, \boldsymbol{\sigma}_s, \boldsymbol{\rho}_k\right)_{s,k}$ remains in $\mathcal{H}$ defined as

$$\mathcal{H} = \Big\{(\mathbf{m}, \mathbf{v}, \mathbf{b}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathbb{R}^{5N} : \text{ for each } i \in \mathcal{N}, \quad 0 \le m_i \le q_{\max}, 0 \le v_i \le q_{\max}^2, \ \underline{b} \le b_i \le \overline{b}, \quad l_i f_{\min} \le \sigma_i \le l_i f_{\max}, \ r_{i,\min} \le \rho_i \le r_{\max} \Big\}. \tag{26}$$

Let $\widehat{\boldsymbol{\Theta}}(t) = \left(\widehat{\mathbf{m}}(t), \widehat{\mathbf{v}}(t), \widehat{\mathbf{b}}(t), \widehat{\boldsymbol{\sigma}}(t), \widehat{\boldsymbol{\rho}}(t)\right) \in \mathcal{H}$ and $\widehat{\boldsymbol{\theta}}_i(t) = (\widehat{m}_i(t), \widehat{b}_i(t))$ for each $i \in \mathcal{N}$, i.e., $\widehat{\boldsymbol{\theta}}_i(t)$ includes only the components in $\widehat{\boldsymbol{\Theta}}(t)$ that affect the quality adaptation of client $i \in \mathcal{N}$. For each client $i \in \mathcal{N}$, we use the variables $\widehat{m}_i(t)$, $\widehat{v}_i(t)$, $\widehat{b}_i(t)$, $\widehat{\sigma}_i(t)$ and $\widehat{\rho}_i(t)$ to track the average dynamics of the parameters $m_{i,s_i}, v_{i,s_i}, b_{i,k}, \sigma_{i,s_i}$ and $\rho_{i,k}$ respectively associated with RAQA (explained in detail in the sequel).

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

Auxiliary differential inclusion related to RAQA

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

$$\dot{\widehat{m}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} \left(E\left[q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right] - \widehat{m}_i(t)\right), \tag{27}$$

$$\dot{\widehat{v}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} \left(E\left[\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right) - \widehat{m}_i(t)\right)^2\right] - \widehat{v}_i(t)\right), \tag{28}$$

$$\dot{\widehat{b}}_i(t) = \frac{1}{\left(1+\overline{\beta}_i\right)} - \frac{l_i}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} + \widehat{z}_i^b\left(\widehat{\boldsymbol{\Theta}}(t)\right), \tag{29}$$

$$\dot{\widehat{\sigma}}_i(t) = \frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} \left(E\left[F_i^\pi\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right)\right] - \widehat{\sigma}_i(t)\right), \tag{30}$$

$$\dot{\widehat{\rho}}_i(t) = \frac{1}{\tau_{slot}} \left(\frac{E\left[r_i^*\left(\widehat{\mathbf{b}}(t), C^\pi\right)\right]}{\tau_{slot}} - \widehat{\rho}_i(t)\right), \tag{31}$$

where

$$u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right) = \tau_{slot} \frac{E\left[l_i F_i^\pi\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right)\right]}{E\left[r_i^*\left(\widehat{\mathbf{b}}(t), C^\pi\right)\right]}, \tag{32}$$

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 $\widehat{z}_i^b\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ mimics the role of the operator $[.]_{\underline{b}}$ in (11), and ensure that $\left(\widehat{\boldsymbol{\Theta}}(t)\right)_{t\ge 0}$ stays in $\mathcal{H}$ (see [21] for a more detailed discussion and see Section 4.3 of [19] for a discussion about projected stochastic approximation). Note that $u_i(.)$ is a set valued map (and hence (27)–(31) describes a differential inclusion) since the denominator $E\left[r_i^*\left(\widehat{\mathbf{b}}(t), C^\pi\right)\right]$ in (32) is a set valued map. Finally, note that the above definition only requires that $\left(\widehat{\boldsymbol{\Theta}}(t)\right)_{t\ge 0}$ is differentiable for *almost* all $t \ge 0$, i.e., we are considering the class of absolutely continuous functions $\left(\widehat{\boldsymbol{\Theta}}(t)\right)_{t\ge 0}$ that satisfy (27)–(31). We can show that the differential inclusion (27)–(31) is well defined, i.e., there exists an absolutely continuous function that solves (27)–(31) for any $\widehat{\boldsymbol{\Theta}}(0) \in \mathcal{H}$. Further, we can show that these solutions are Lipschitz continuous and stay in $\mathcal{H}$ and hence are bounded.

By comparing (27)–(31) against RAQA update rules (9)–(11) and (23)–(25), we see that the differential inclusion (27)–(31) reflects the average dynamics of the evolution of parameters in RAQA. For instance, this is apparent when we compare the update rule (27) against (10). Note that the rate of change of $\widehat{m}_i(t)$ given in (27) has a scaling term $\frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}$ which corresponds to the segment download rate of client i at time t and $u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)$ defined in (32) is expected segment download duration of client i at time t. This scaling is naturally expected since $\widehat{m}_i(t)$, $\widehat{v}_i(t)$, and $\widehat{\sigma}_i(t)$ correspond to RAQA parameters that are updated when a segment download is completed, and thus we can view $\frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}$ as the update rate associated with these parameters. Similarly, we can view the constant scaling term $\frac{1}{\tau_{slot}}$ in (31) describing the evolution of $\widehat{\rho}_i(t)$ as the corresponding update rate by noting that the associated (auxiliary) RAQA parameter $\rho_{i,k}$ is updated at the beginning of every slot, i.e., once every τ_{slot} seconds. Finally, note that Eq. (29) describing the evolution of $\widehat{b}_i(t)$ can be rewritten as

$$\dot{\widehat{b}}_i(t) = \frac{1}{\tau_{slot}} \left(\frac{\tau_{slot}}{\left(1+\overline{\beta}_i\right)}\right) - \frac{l_i}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)} + \widehat{z}_i^b\left(\widehat{\boldsymbol{\Theta}}(t)\right),$$

and presence of the two scaling terms $\frac{1}{\tau_{slot}}$ and $\frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}$ reflects the fact that the corresponding RAQA parameter $b_{i,k}$ is updated at the beginning of every slot (using (9)) and when a segment download of client i is completed (using (11)). Thus, we can expect that (27)–(31) captures the average dynamics of RAQA, and the presence of the client dependent update rates $\left(\frac{1}{u_i\left(\widehat{\boldsymbol{\Theta}}(t)\right)}\right)_{i\in\mathcal{N}}$ reflects the *asynchronous* nature of the evolution of RAQA parameters where different clients are updating their parameters at their own (possibly time varying) rates.

Next, we define certain classes of policies.

Definition 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 each slot k when $C_k = c$, based only on the allocation constraint in that slot.

Definition 2. Feasible stationary resource allocation policy: A stationary resource allocation policy $\left((\mathbf{r}(c))_{c\in\mathcal{C}}\right)$ is feasible if

$\mathbf{r}(c) \ge \mathbf{r}_{\min}$ and $c(\mathbf{r}(c)) \le 0, \ \forall\, c \in \mathcal{C}$.

Definition 3. Stationary quality adaptation policy for 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 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 client i that chooses quality $q_i(f_i)$ for each segment s with QR trade-off f_i, based only on the QR trade-off of that segment.

Definition 4. Feasible stationary quality adaptation policy for client i: We say that a stationary quality adaptation policy $(q_i(f_i))_{f_i\in\mathcal{F}_i}$ for client i is feasible if $0 \le q_i(f_i) \le q_{\max}$ for each $f_i \in \mathcal{F}_i$.

Next, we define the set $\widetilde{\mathcal{H}} \subset \mathbb{R}^{5N}$ as

$$\widetilde{\mathcal{H}} = \Bigg\{ (\mathbf{m}, \mathbf{v}, \mathbf{b}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathcal{H} : \exists \text{ a feasible stationary resource allocation policy } (\mathbf{r}(c))_{c\in\mathcal{C}} \text{ s.t. } \frac{\mathbb{E}\left[r_i\left(C^\pi\right)\right]}{\tau_{slot}} = \rho_i \ \forall\, i \in \mathcal{N}; \text{ for each } i \in \mathcal{N}, \ \exists \text{ there is a feasible stationary quality adaptation scheme } \left((q_i(f_i))_{f_i\in\mathcal{F}_i}\right) \text{ such that } \mathbb{E}\left[q_i\left(F_i^\pi\right)\right] = m_i, \operatorname{Var}\left(q_i\left(F_i^\pi\right)\right) \le v_i \le q_{\max}^2, \ \mathbb{E}\left[F_i^\pi\left(q_i\left(F_i^\pi\right)\right)\right] \le \sigma_i \le f_{\max} \Bigg\}. \tag{33}$$

We can view $\widetilde{\mathcal{H}}$ as the set of 'achievable' parameters in $\mathcal{H}$, i.e., for $(\mathbf{m}, \mathbf{v}, \mathbf{b}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathcal{H}$ 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 mean quality of m_i, a variance in quality which is at least v_i and mean segment size which is at least σ_i.

It can be verified that $\widetilde{\mathcal{H}}$ is a bounded, closed and convex set (using an approach similar to Lemma 5 (b) in [18]). Hence, we conclude that for any $\Theta \in \mathcal{H}$, there exists a unique projection of $\widetilde{\Theta} \in \mathcal{H}$ onto the set $\widetilde{\mathcal{H}}$. Let $\tilde{}$ denote this projection operator. 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 5. *There exists a finite constant $\chi_0 > 0$ such that for any initialization $\widehat{\Theta}(0) \in \mathcal{H}$,*

$$\frac{d}{dt} d_{5N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right) \leq -\chi_0 d_{5N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right).$$

Hence, $\lim_{t\to\infty} d_{5N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right) = 0$.

Proof. The result is an application of a generalization of Lemma 3 in [26]. ■

The next result provides the main convergence result for the differential inclusion (27)–(31) which states that $\widehat{\Theta}(t)$ converges to the following set

$$\mathcal{H}^* = \Big\{(\mathbf{m}, \mathbf{v}, \mathbf{b}, \boldsymbol{\sigma}, \boldsymbol{\rho}) \in \mathcal{H} : \left(\boldsymbol{\rho}, \left(h_i^B(b_i)\right)_{i\in\mathcal{N}}\right) \in \mathcal{X}^\pi, \text{ and for each } i \in \mathcal{N},\ m_i = m_i^\pi,\ v_i = v_i^\pi \Big\} \quad (34)$$

Observe that (based on (34), (26) and (33))

$$\mathcal{H}^* \subset \widetilde{\mathcal{H}} \subset \mathcal{H}.$$

Recall that Theorem 2 suggested that we can prove Theorem 1, if we can show that the updates (10)–(11) guide RAQA parameters $\left(\boldsymbol{\theta}_{i,s}\right)_{s\geq 1}$ of client i to the set $\mathcal{H}_i^*$ (defined in (22)) for each 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 client i's components and 'dropping' the components $(\mathbf{v}, \boldsymbol{\sigma}, \boldsymbol{\rho})$). Hence, the following result along with Theorem 4 (which relates evolution of RAQA parameters to the differential inclusion) help us to establish the desired convergence property for RAQA parameters.

Theorem 3 *(a). For $\widehat{\Theta} = \left(\widehat{\mathbf{m}}, \widehat{\mathbf{v}}, \widehat{\mathbf{b}}, \widehat{\boldsymbol{\sigma}}, \widehat{\boldsymbol{\rho}}\right) \in \mathcal{H}$, and some $(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi) \in \mathcal{X}^\pi$, let*

$$L\left(\widehat{\Theta}\right) := -\sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right) l_i\left(\widehat{m}_i - \eta_i \widehat{v}_i\right) \quad (35)$$

$$+\sum_{i\in\mathcal{N}} \left(l_i b_i^\pi \widehat{\sigma}_i - \tau_{slot} b_i^\pi \widehat{\rho}_i\right) + \sum_{i\in\mathcal{N}} \sigma_i^\pi \int_{\underline{b}}^{\widehat{b}_i} \left(h_i^B(e) - b_i^\pi\right) de$$

$$+\sum_{i\in\mathcal{N}} \left(1+\overline{\beta}_i\right) l_i \left(\widehat{m}_i - m_i^\pi\right)^2 + \frac{\chi_2}{\chi_0} d\left(\widehat{\Theta}, \widetilde{\mathcal{H}}\right),$$

where χ_0 is the positive constant from Lemma 5, and χ_2 is an appropriately chosen (large) positive constant. If $\widehat{\Theta}(0) \in \mathcal{H}$, then for almost all t

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

(b) If $\widehat{\Theta}(0) \in \mathcal{H}$, then $\lim_{t\to\infty} d_{5N}\left(\widehat{\Theta}(t), \mathcal{H}^*\right) = 0$.

Proof of the above theorem is given in Appendix.

5.3. RAQA parameters also converge to optimal parameter set

The main focus of this subsection is Theorem 4 which relates RAQA to the auxiliary differential inclusion (27)–(31), and obtains the desired convergence result for RAQA by using the convergence result in Theorem 3 for the differential inclusion. Our approach here relies on viewing the update equations ((9)–(11) and (23)–(25)) of RAQA as an asynchronous stochastic approximation update equation (see Chapter 12 of [19] for a detailed discussion on asynchronous stochastic approximation) to relate RAQA to the differential inclusion using tools from the theory of stochastic approximation.

Next, we define two auxiliary variables $b_{R,i,k}$ and b_{Q,i,s_i+1}. At the beginning of slot k, let $b_{R,i,k} = b_{i,k}$ for each $i \in \mathcal{N}$ and thus the variable stores the value of $b_{i,k}$ used while deciding allocation for kth slot. In slot k, if any client $i \in \mathcal{N}$ finishes download of s_i th segment, let $b_{Q,i,s_i+1} = b_{i,k+1}$, and thus the variable stores the value of $b_{i,k}$ used while deciding the quality for client i's (s_i+1)-th segment. In the following, we use superscript ϵ on RAQA parameters $(m_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(v_{i,s}^\epsilon)_{i\in\mathcal{N}}$, $(b_{Q,i,s}^\epsilon)_{i\in\mathcal{N}}$, $(b_{R,i,k}^\epsilon)_{i\in\mathcal{N}}$, $(b_{i,k}^\epsilon)_{i\in\mathcal{N}}$, $(\sigma_{i,s}^\epsilon)_{i\in\mathcal{N}}$ and $(\rho_{i,k}^\epsilon)_{i\in\mathcal{N}}$ to emphasize their dependence on ϵ (see RAQA updates in (9)–(11) to see the dependence). We refer to the update of RAQA parameters $\left(m_{i,s_i}, b_{i,k}\right)$ in (10)–(11) carried out after the selection of segment quality for client i (following a segment download) as a Q_i-update. Let $\delta\tau_{Q,i,s}^\epsilon$ denote the time (in seconds) between the sth and $(s+1)$th Q_i-updates. Let $\tau_{Q,i,s}^\epsilon = \epsilon\sum_{j=0}^{s-1}\delta\tau_{Q,i,j}^\epsilon$. denote ϵ times the cumulative time for the first s Q_i-updates.

Next, we define time interpolated processes $\widehat{\Theta}^\epsilon(t) = \left(\widehat{\mathbf{m}}^\epsilon(t), \widehat{\mathbf{v}}^\epsilon(t), \widehat{\mathbf{b}}^\epsilon(t), \widehat{\boldsymbol{\sigma}}^\epsilon(t), \widehat{\boldsymbol{\rho}}^\epsilon(t)\right)$ associated with RAQA's parameters. For each $i \in \mathcal{N}$ and for $t \in \left[\tau_{Q,i,s}^\epsilon, \tau_{Q,i,s+1}^\epsilon\right)$, let $\widehat{m}_i^\epsilon(t) = m_{i,s}^\epsilon$, $\widehat{v}_i^\epsilon(t) = v_{i,s}^\epsilon$, $\widehat{b}_{Q,i}^\epsilon(t) = b_{Q,i,s}^\epsilon$, and $\widehat{\sigma}_i^\epsilon(t) = \sigma_{i,s}^\epsilon$. Also, for $t \in [k\tau_{slot}\epsilon, (k+1)\tau_{slot}\epsilon)$, let $\widehat{b}_{R,i}^\epsilon(t) = b_{R,i,k}^\epsilon$ and $\widehat{\rho}_i^\epsilon(t) = \rho_{i,k}^\epsilon$. For each t, let

$$\widehat{\Theta}_Q^\epsilon(t) = \left(\widehat{\mathbf{m}}^\epsilon(t), \widehat{\mathbf{v}}^\epsilon(t), \widehat{\mathbf{b}}_Q^\epsilon(t), \widehat{\boldsymbol{\sigma}}^\epsilon(t), \widehat{\boldsymbol{\rho}}^\epsilon(t)\right),$$
$$\widehat{\Theta}_R^\epsilon(t) = \left(\widehat{\mathbf{m}}^\epsilon(t), \widehat{\mathbf{v}}^\epsilon(t), \widehat{\mathbf{b}}_R^\epsilon(t), \widehat{\boldsymbol{\sigma}}^\epsilon(t), \widehat{\boldsymbol{\rho}}^\epsilon(t)\right),$$

The next result states that for small enough ϵ, the time interpolated versions of RAQA parameters $\widehat{\Theta}_Q^\epsilon(.)$ and $\widehat{\Theta}_R^\epsilon(.)$ stay close to the set $\mathcal{H}^*$ (defined in (34)) most of the time over long time windows.

Theorem 4. *Let $\widehat{\Theta}_Q^\epsilon(0) = \widehat{\Theta}^\epsilon(0) \in \mathcal{H}$. Then, the fraction of time in the time interval $[0, T]$ that $\widehat{\Theta}_Q^\epsilon(.)$ and $\widehat{\Theta}_R^\epsilon(.)$ spend in a small neighborhood of $\mathcal{H}^*$ converges to one in probability as $\epsilon \to 0$ and $T \to \infty$.*

Proof sketch. This result follows from an extension of Theorem 3.4 in Chapter 12 of [19] which relates asynchronous stochastic approximation (10)–(11) to its associated differential inclusion (27)–(31), and using Theorem 3 (regarding convergence of differential inclusion to $\mathcal{H}^*$). Theorem 3.4 cannot be directly applied mainly because condition (A3.8) (given in SEction 12.3.3, page 418 of [19]) concerning the time between the (asynchronous) updates is not be satisfied in our problem setting (discussed later in the proof). The details of extension are omitted for brevity and are discussed in [21]. ■

We have the following corollary of Theorem 4 which says that for small enough ϵ and after running RAQA for long enough, client i's RAQA parameter stays close to $\mathcal{H}_i^*$(defined in (22)) most of the time with high probability. The corollary can be proved using Theorem 4 noting that the amount of time between updates is bounded.

Corollary 1. *Let $\widehat{\Theta}^{\epsilon}(0) \in \mathcal{H}$ 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 $(\boldsymbol{\theta}_{i,s})_{1\le s\le S_\epsilon}$ is in a δ-neighborhood of $\mathcal{H}_i^*$ converges to one in probability as $\epsilon \to 0$ and $S \to \infty$.*

5.4. Proof of Theorem 1

We have now obtained all the intermediate results required to prove Theorem 1 which is given below.

Proof of Theorem 1. To show part (a) (i.e., (12)), let $T_i(S)$ (measured in seconds) denote the time at which the download of first S segments of client i completes. From (9) and (11), we get below lower bound on $b_{Q,i,S}$:

$$b_{Q,i,S} \ge b_{Q,i,0} + \epsilon\left(\frac{\tau_{slot}\left\lfloor\frac{T_i(S)}{\tau_{slot}}\right\rfloor}{(1+\overline{\beta}_i)} - l_i S\right)$$

$$\ge b_{Q,i,0} - \frac{\epsilon\tau_{slot}}{(1+\overline{\beta}_i)} + \epsilon\left(\frac{T_i(S)}{(1+\overline{\beta}_i)} - l_i S\right).$$

Hence,

$$\frac{T_i(S)}{lS} \le (1+\overline{\beta}_i)\left(1+\left(\frac{b_{Q,i,S}-b_{Q,i,0}+\frac{\epsilon\tau_{slot}}{(1+\overline{\beta}_i)}}{\epsilon l_i S}\right)\right). \quad (36)$$

Now, if we let $K_i(S)$ denote the (random variable associated with) the number of slots which client i takes to download S segments, then we can express the term appearing in the left hand side of above inequality as

$$\frac{T_i(S)}{l_i S} = \frac{\tau_{slot}\frac{\sum_{s=1}^{S} l_i f_{i,s}\left(q_{i,s}^*\right)}{\frac{1}{K_i(S)}\sum_{k=1}^{K_i(S)} r_{i,k}^*}}{lS} + o(S). \quad (37)$$

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 (4), (37), (36) and the fact that $b_{Q,i,S}$ is bounded (see Lemma 1), we can conclude that (12) also holds.

Next, we prove part (b) of Theorem 1 regarding the optimality of RAQA. Using Corollary 1 (which says that $(\boldsymbol{\theta}_{i,s})_{1\le s\le S_\epsilon}$ essentially converges to $\mathcal{H}_i^*$) and Lemma 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 $\boldsymbol{\theta}_i^\pi \in \mathcal{H}_i^*$

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

goes to zero in probability. Now, part (b) of Theorem 1 follows from the above observation and Theorem 2. ■

6. Conclusions and future directions

We developed a simple online algorithm for optimizing asynchronous streaming of bufferable information flows, well suited for advanced wireless networks carrying DASH-based video and audio, and with potential applications in streaming of AR/VR. Several interesting extensions (e.g., more general QoE models, cost constraints, discrete quality-rate tradeoffs) of the settings considered in this paper and extensive evaluation using simulations can be found in [21] where the focus is on DASH-based video streaming.

Interesting future directions might include exploring the potential for *learning* user preferences, and developing 'RAQA-like' algorithms for networks with contention based medium access potentially by modulating the back-off timers using information about parameters like $b_{i,k}$.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Prof Gustavo de Veciana reports financial support was provided by Cisco Systems Inc. Prof Gustavo de Veciana reports financial support was provided by Intel.

Data availability

No data was used for the research described in the article.

Acknowledgments

We fondly acknowledge discussions and suggestions from Prof. Ari Arapostathis that paved the way to this work. This research was supported in part by Intel and Cisco under the VAWN program, and by the NSF under Grant CNS-1910112. This work was supported in part by the Department of Science and Technology, Government of India through the FIST Scheme under the Grant SR/FST/ET-I/2017/68.

Appendix. Proof of Theorem 3

Proof. The proof of part (a) involves analysis of drift of the Lyapunov function $L(.)$. The choice of several terms in $L(.)$ is motivated by Lyapunov functions in [13,20], and has many novel elements. For instance, the term $\sum_{i\in\mathcal{N}}(1+\overline{\beta}_i)\, l_i\left(\widehat{m}_i - m_i^\pi\right)^2$ allows us to accommodate objectives involving variability terms and the terms in $\sum_{i\in\mathcal{N}} b_i^\pi\left(l_i\widehat{\sigma}_i - \tau_{slot}\widehat{\rho}_i\right) + \sum_{i\in\mathcal{N}}\sigma_i^\pi\int_{\underline{b}}^{\widehat{b}_i}\left(h_i^B(e) - b_i^\pi\right)de$ which allow us to accommodate the rebuffering constraints. Further, our convergence result is for a differential inclusion associated with an algorithm RAQA which, unlike those in [13,20], uses *asynchronous* updates. To just simplify the notation used in this proof, we assume $\overline{\beta}_i = 0$ and $\eta_i = 1$.

Since $(\boldsymbol{\rho}^\pi, \mathbf{b}^\pi) \in \mathcal{X}^\pi$, 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 $\mathbf{b}^\pi$ as the associated optimal Lagrange multipliers for the constraints (17) respectively.

Using the definition of $L(.)$ and (27)–(31), we have that

$$\frac{dL\left(\widehat{\Theta}(t)\right)}{dt} \le -\sum_{i\in\mathcal{N}}\frac{l_i}{u_i(t)} \quad (38)$$
$$\left(\left(E\left[q_i^*(t)\right] - \widehat{m}_i(t)\right) - \left(E\left[\left(q_i^*(t) - \widehat{m}_i(t)\right)^2\right] - \widehat{v}_i(t)\right)\right)$$
$$+\sum_{i\in\mathcal{N}}\frac{2l_i}{u_i(t)}\left(\widehat{m}_i(t) - m_i^\pi\right)\left(E\left[q_i^*(t)\right] - \widehat{m}_i(t)\right)$$
$$+\sum_{i\in\mathcal{N}}\frac{l_i b_i^\pi}{u_i(t)}\left(E\left[F_i^\pi\left(q_i^*(t)\right)\right] - \widehat{\sigma}_i(t)\right)$$
$$-\sum_{i\in\mathcal{N}} b_i^\pi\left(\frac{E\left[r_i^*(t)\right]}{\tau_{slot}} - \widehat{\rho}_i(t)\right) - \chi_2 d_{5N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right)$$
$$+\sum_{i\in\mathcal{N}}\sigma_i^\pi\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(1 - \frac{l_i}{u_i(t)}\right)$$
$$+\sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{E\left[l_i F_i^\pi\left(q_i^*(t)\right)\right]}{u_i(t)} - \frac{E\left[r_i^*(t)\right]}{\tau_{slot}}\right).$$

For brevity, we have not explicitly indicated the dependence of many terms above on $\widehat{\Theta}(t)$. For instance, $u_i(t)$ is shorthand for $u_i\left(\widehat{\Theta}(t)\right)$, and $E\left[l_i q_i^*(t)\right]$ is shorthand for $E\left[l_i q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right)\right]$ where $\widehat{\boldsymbol{\theta}}_i(t) = (\widehat{m}_i(t), \widehat{b}_i(t))$. Also, note that the last term is zero-valued (see (32)). Term involving $\widehat{z}_i^b(\widehat{\Theta}(t))$ has also been dropped from right hand side of (38) as it is non-positive. To see this, note that $\widehat{z}_i^b(\widehat{\Theta}(t)) \geq 0$ which is equal to zero unless $\widehat{b}_i(t) = \underline{b}$ (from Eq. (29)) for which $h_i^B(\widehat{b}_i(t)) = 0 \leq b_i^\pi$. Hence, $\sum_{i\in\mathcal{N}} \sigma_i^\pi \left(h_i^B(\widehat{b}_i(t)) - b_i^\pi\right) \widehat{z}_i^b(\widehat{\Theta}(t)) \leq 0$.

Consider the right hand side of (38), and group the terms containing $q_i^*(t)$ except those in the line with the term $\left(\widehat{m}_i(t) - m_i^\pi\right)$ to note that we have negative of a scaled (by $1/u_i(t)$) version of the expectation of the objective of QA$_i$ $\left(\widehat{\boldsymbol{\theta}}_i(t), f_i\right)$, 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 $\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t) = (m_i^\pi, \widehat{b}_i(t))$, i.e., $\widehat{\boldsymbol{\theta}}_i(t)$ with the first component set to m_i^π (defined in (18)). 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 $\Delta_1\left(\widehat{\Theta}(t)\right)$, and rewrite (38) as

$$
\begin{aligned}
\frac{dL\left(\widehat{\Theta}(t)\right)}{dt} &\leq \Delta_1\left(\widehat{\Theta}(t)\right) \\
&- \sum_{i\in\mathcal{N}} l_i \left(\frac{1}{u_i(t)}\left(E\left[q_i^{*,m_i^\pi}(t)\right] - \widehat{m}_i(t)\right) - \frac{1}{u_i(t)}\left(E\left[\left(q_i^{*,m_i^\pi}(t) - m_i^\pi\right)^2\right] - \widehat{v}_i(t)\right)\right) \\
&+ \sum_{i\in\mathcal{N}} \frac{l_i b_i^\pi}{u_i(t)}\left(E\left[F_i^\pi\left(q_i^{*,m_i^\pi}(t)\right)\right] - \widehat{\sigma}_i(t)\right) \\
&- \sum_{i\in\mathcal{N}} b_i^\pi \left(\frac{E\left[r_i^*(t)\right]}{\tau_{slot}} - \widehat{\rho}_i(t)\right) - \chi_2 d_{5N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right) \\
&+ \sum_{i\in\mathcal{N}} \sigma_i^\pi \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(1 - \frac{l}{u_i(t)}\right) \\
&+ \sum_{i\in\mathcal{N}} \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right) \left(\frac{E\left[l_i F_i^\pi\left(q_i^{*,m_i^\pi}(t)\right)\right]}{u_i(t)} - \frac{E\left[r_i^*(t)\right]}{\tau_{slot}}\right),
\end{aligned}
\tag{39}
$$

where

$$
\begin{aligned}
&\Delta_1\left(\widehat{\Theta}(t)\right) \\
&= -\sum_{i\in\mathcal{N}} \frac{l_i}{u_i(t)} E\Big[\Big(\phi^Q\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right), \widehat{\boldsymbol{\theta}}_i(t), F_i^\pi\right) \\
&\quad - \phi^Q\left(q_i^*\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), F_i^\pi\right), \widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), F_i^\pi\right) \\
&\quad - 2\left(\widehat{m}_i(t) - m_i^\pi\right)\left(q_i^*(t) - \widehat{m}_i(t)\right)\Big)\Big].
\end{aligned}
\tag{40}
$$

From the definition (32) of $u_i(t)$, we have that for each $i \in \mathcal{N}$

$$
u_{\min} := \frac{\tau_{slot} l f_{\min}}{r_{\max}}, \quad u_{\max} := \frac{\tau_{slot} l f_{\max}}{r_{\min}} \tag{41}
$$

are lower and upper bounds respectively on $u_i(t)$.

If we group the terms containing $q_i^{*,m_i^\pi}(t)$ and $\mathbf{r}^*(t)$, we find that the right hand side of (39) contains negative of scaled versions of optimal value of objective functions of QA$_i$ $\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$ (i.e., $\phi^Q\left(q_i^*(t), \widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$) and those of RA$\left(\widehat{\mathbf{b}}(t), c\right)$ (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 QA$_i$ $\left(\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}(t), f_i\right)$ and RA$\left(\widehat{\mathbf{b}}(t), c\right)$, 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 (39) (obtained by replacing $q_i^{*,m_i^\pi}(t)$ and $\mathbf{r}^*(t)$ with $q_i^\pi(f)$ and $\mathbf{r}^\pi(c)$ in (39) and adding the correction term $\Delta_2\left(\widehat{\Theta}(t)\right)$ associated with this replacement)

$$
\begin{aligned}
\frac{dL\left(\widehat{\Theta}(t)\right)}{dt} &\leq \Delta_1\left(\widehat{\Theta}(t)\right) + \Delta_2\left(\widehat{\Theta}(t)\right) \\
&- \sum_{i\in\mathcal{N}} l_i \left(\frac{1}{u_i(t)}\left(m_i^\pi - \widetilde{\widehat{m}_i(t)}\right) - \frac{1}{u_i(t)}\left(E\left[\left(q_i^\pi\left(F_i^\pi\right) - m_i^\pi\right)^2\right] - \widetilde{\widehat{v}_i(t)}\right)\right) \\
&+ \sum_{i\in\mathcal{N}} \frac{l_i b_i^\pi}{u_i(t)}\left(\sigma_i^\pi - \widetilde{\widehat{\sigma}_i(t)}\right) - \sum_{i\in\mathcal{N}} b_i^\pi \left(\rho_i^\pi - \widetilde{\widehat{\rho}_i(t)}\right) \\
&+ \sum_{i\in\mathcal{N}} \sigma_i^\pi \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(1 - \frac{l_i}{u_i(t)}\right) \\
&+ \sum_{i\in\mathcal{N}} \left(h_i^B\left(\widehat{b}_i(t)\right) - b_i^\pi\right)\left(\frac{\sigma_i^\pi l_i}{u_i(t)} - \rho_i^\pi\right) \\
&- \chi_2 d_{8N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right) + l_2\left(\widehat{\Theta}(t)\right),
\end{aligned}
\tag{42}
$$

where m_i^π, v_i^π and σ_i^π are defined in (18)–(20), (and $\boldsymbol{\rho}^\pi$ was chosen at the beginning of the proof — see below (36))

$$
\begin{aligned}
\Delta_2\left(\widehat{\Theta}(t)\right) &= -\frac{1}{\tau_{slot}} E\Big[\phi^R\left(\mathbf{r}^*(t), \widehat{\mathbf{b}}(t), C^\pi\right) \\
&\qquad - \phi^R\left(\mathbf{r}^\pi\left(C^\pi\right), \widehat{\mathbf{b}}(t), C^\pi\right)\Big] \\
&- \sum_{i\in\mathcal{N}} \frac{l_i}{u_i(t)} E\Big[\Big(\phi^Q\left(q_i^*\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), F_i^\pi\right), \widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), F_i^\pi\right) \\
&\qquad - \phi^Q\left(q_i^\pi\left(F_i^\pi\right), \widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t), F_i^\pi\right)\Big)\Big],
\end{aligned}
\tag{43}
$$

and $\widetilde{\cdot}$ is the projection of elements in $\mathcal{H}$ to the set $\widetilde{\mathcal{H}}$. Here $\left(\widetilde{\widehat{\mathbf{m}}(t)}, \widetilde{\widehat{\mathbf{v}}(t)}, \widetilde{\widehat{\mathbf{b}}(t)}, \widetilde{\widehat{\boldsymbol{\sigma}}(t)}, \widetilde{\widehat{\boldsymbol{\rho}}(t)}\right) := \widetilde{\widehat{\Theta}(t)}$.

Due to the definition of $\widetilde{\mathcal{H}}$ (see (33)), $\widetilde{\widehat{\mathbf{b}}(t)} = \widehat{\mathbf{b}}(t)$. Also, for each $i \in \mathcal{N}$,

$$
\widetilde{\widehat{\boldsymbol{\theta}}_i(t)} := \left(\widetilde{\widehat{m}_i(t)}, \widehat{b}_i(t)\right), \widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^\pi)}}(t) := \left(m_i^\pi, \widehat{b}_i(t)\right).
$$

Note that we have replaced components of $\widehat{\mathbf{m}}(t)$, $\widehat{\mathbf{v}}(t)$, $\widehat{\boldsymbol{\sigma}}(t)$ and $\widehat{\boldsymbol{\rho}}(t)$ appearing in (39) with those of $\widetilde{\widehat{\mathbf{m}}(t)}$, $\widetilde{\widehat{\mathbf{v}}(t)}$, $\widetilde{\widehat{\boldsymbol{\sigma}}(t)}$ and $\widetilde{\widehat{\boldsymbol{\rho}}(t)}$ respectively, and in (42), we have added the function $l_2(.)$ defined below to account for these replacements:

$$
\begin{aligned}
l_2\left(\widehat{\Theta}(t)\right) &= -\sum_{i\in\mathcal{N}} l_i \left(\left(\widetilde{\widehat{m}_i(t)} - \widehat{m}_i(t)\right) - \frac{1}{u_i(t)}\left(\widetilde{\widehat{v}_i(t)} - \widehat{v}_i(t)\right)\right) \\
&- \sum_{i\in\mathcal{N}} \frac{l_i b_i^\pi}{u_i(t)}\left(\widetilde{\widehat{\sigma}_i(t)} - \widehat{\sigma}_i(t)\right) - \sum_{i\in\mathcal{N}} b_i^\pi \left(\widetilde{\widehat{\rho}_i(t)} - \widehat{\rho}_i(t)\right).
\end{aligned}
\tag{44}
$$

Since all terms above are bounded, there exists some large enough finite constant χ_4 such that

$$
l_2\left(\widehat{\Theta}(t)\right) \leq \chi_4 d_{8N}\left(\widehat{\Theta}(t), \widetilde{\mathcal{H}}\right) \tag{45}
$$

holds for any $\widehat{\Theta}(t) \in \mathcal{H}$. Thus, we can use the observations in (45) along with (42) to conclude that

$$
\begin{aligned}
\frac{dL\left(\widehat{\Theta}(t)\right)}{dt} &\leq \Delta_1\left(\widehat{\Theta}(t)\right) + \Delta_2\left(\widehat{\Theta}(t)\right) \\
&- \sum_{i\in\mathcal{N}} l_i \left(\frac{1}{u_i(t)}\left(m_i^\pi - \widetilde{\widehat{m}_i(t)}\right)\right.
\end{aligned}
\tag{46}
$$

$$-\frac{1}{u_i(t)}\left(E\left[\left(q_i^{\pi}\left(F_i^{\pi}\right)-m_i^{\pi}\right)^2\right]-\widetilde{v_i(t)}\right)\right)$$
$$+\sum_{i\in\mathcal{N}}\frac{l_i b_i^{\pi}}{u_i(t)}\left(\sigma_i^{\pi}-\widetilde{\sigma_i(t)}\right)-\sum_{i\in\mathcal{N}}b_i^{\pi}\left(\rho_i^{\pi}-\widetilde{\rho_i(t)}\right)$$
$$+\sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right)-b_i^{\pi}\right)\left(\sigma_i^{\pi}-\rho_i^{\pi}\right)$$
$$+\left(\chi_3+\chi_4+\chi_5-\chi_2\right)d_{8N}\left(\widehat{\boldsymbol{\Theta}}(t),\widetilde{\mathcal{H}}\right).$$

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

$$\Delta(\boldsymbol{\Theta})=\Delta_1(\boldsymbol{\Theta})+\Delta_2(\boldsymbol{\Theta})+\Delta^{(b)}(\boldsymbol{\Theta})+\Delta^{(\pi,q)}(\boldsymbol{\Theta})+\Delta^{(\pi,r)}(\boldsymbol{\Theta})+\Delta_3(\boldsymbol{\Theta}), \quad (47)$$

where

$$\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)=-\sum_{i\in\mathcal{N}}\frac{l_i}{u_i(t)}\left(\sum_{i\in\mathcal{N}}b_i^{\pi}\left(\sigma_i^{\pi}-\widetilde{\sigma_i(t)}\right)+\left(\left(m_i^{\pi}-\widetilde{m_i(t)}\right)-\left(v_i^{\pi}-\widetilde{v_i(t)}\right)\right)\right),$$
$$\Delta^{(\pi,r)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)=-\sum_{i\in\mathcal{N}}b_i^{\pi}\left(\rho_i^{\pi}-\widetilde{\rho_i(t)}\right),$$
$$\Delta^{(b)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)=\sum_{i\in\mathcal{N}}\left(h_i^B\left(\widehat{b}_i(t)\right)-b_i^{\pi}\right)\left(\sigma_i^{\pi}-\rho_i^{\pi}\right),$$
$$\Delta_3\left(\widehat{\boldsymbol{\Theta}}(t)\right)=-d_{5N}\left(\widehat{\boldsymbol{\Theta}}(t),\widetilde{\mathcal{H}}\right).$$

Hence, we can rewrite (46) as follows:

$$\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt}\leq\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right). \quad (48)$$

Next, we show that all the functions $\Delta_1(\boldsymbol{\Theta})$, $\Delta_2(\boldsymbol{\Theta})$, $\Delta^{(b)}(\boldsymbol{\Theta})$, $\Delta^{(\pi,q)}(\boldsymbol{\Theta})$, $\Delta^{(\pi,r)}(\boldsymbol{\Theta})$ and $\Delta_3(\boldsymbol{\Theta})$, 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 $\boldsymbol{\Theta}\notin\mathcal{H}^*$.

To show non-positivity of $\Delta^{(b)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$, note that complementary slackness conditions for (17) in OPTSTAT implies that $\sigma_i^{\pi}=\frac{\rho_i^{\pi}}{\tau_{slot}}$ if $b_i^{\pi}>0$. Further, feasibility of the optimal solution (with respect to (17)) implies that $\sigma_i^{\pi}\leq\frac{\rho_i^{\pi}}{\tau_{slot}}$ $i\in\mathcal{N}$.

Next, consider $\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)$. Since $\widetilde{\widehat{\boldsymbol{\Theta}}(t)}\in\widetilde{\mathcal{H}}$, and $\left((q_i^{\pi}(f_i))_{f_i\in\mathcal{F}_i}\right)$ is the unique optimal solution (using Lemma 4(a)), we can show that $\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)\leq 0$ and

$$\Delta^{(\pi,q)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)=0 \text{ only if } \widetilde{\widehat{\mathbf{m}}(t)}=\mathbf{m}^{\pi} \text{ and } \widetilde{\widehat{\mathbf{v}}(t)}=\mathbf{v}^{\pi}. \quad (49)$$

Using similar arguments along with Lemma 4(b), we can show that $\Delta^{(\pi,r)}\left(\widehat{\boldsymbol{\Theta}}(t)\right)\leq 0$.

Next, consider $\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right)$. Using Lemma 3(c), we can show that $\Delta_1\left(\widehat{\boldsymbol{\Theta}}(t)\right)\leq 0$. Next, consider $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right)$. Since $q_i^*\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^{\pi})}}(t),f_i\right)$ and $\mathbf{r}^*(t)$ are optimal solutions to $\text{QA}\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^{\pi})}}(t),f_i\right)$ and $\text{RA}\left(\widehat{\mathbf{b}}(t),c\right)$ respectively, $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right)\leq 0$.

Also, $\Delta_3(\boldsymbol{\Theta})=-d_{5N}\left(\widehat{\boldsymbol{\Theta}},\widetilde{\mathcal{H}}\right)$ is non-positive, and

$$\Delta_3\left(\widehat{\boldsymbol{\Theta}}(t)\right)=0 \text{ only if } \widetilde{\widehat{\mathbf{m}}(t)}=\widehat{\mathbf{m}}(t),\ \widetilde{\widehat{\mathbf{v}}(t)}=\widehat{\mathbf{v}}(t) \text{ and } \widetilde{\widehat{\boldsymbol{\rho}}(t)}=\widehat{\boldsymbol{\rho}}(t). \quad (50)$$

Next, we argue that $\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right)=0$ only if $\left(\widehat{\boldsymbol{\rho}}(t),\left(h_i^B\left(\widehat{b}_i(t)\right)\right)_{i\in\mathcal{N}}\right)\in\mathcal{X}^{\pi}$. 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 (49) and (50), we can conclude that $\widehat{\mathbf{m}}(t)=\mathbf{m}^{\pi}$ and $\widehat{\mathbf{v}}(t)=\mathbf{v}^{\pi}$. We also have that $\Delta_2\left(\widehat{\boldsymbol{\Theta}}(t)\right)=0$, and hence

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

where recall that $\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^{\pi})}}(t)=\left(m_i^{\pi},\widehat{b}_i(t)\right)$. Hence, $\mathbf{r}^{\pi}(c)$ is an optimal solution to $\text{RA}\left(\widehat{\mathbf{b}}(t),c\right)$ for each $c\in\mathcal{C}$, and $q_i^{\pi}(f_i)$ is an optimal solution to $\text{QA}\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^{\pi})}}(t),f_i\right)$ for each $f_i\in\mathcal{F}_i$ and $i\in\mathcal{N}$. Now, we can use KKT conditions and associated optimal Lagrange multipliers for $\text{QA}\left(\widetilde{\widehat{\boldsymbol{\theta}}_i^{(m_i^{\pi})}}(t),f_i\right)$ and $\text{RA}\left(\widehat{\mathbf{b}}(t),c\right)$ to build optimal Lagrange multipliers including $\boldsymbol{b}^{\pi}$ satisfying KKT conditions for OPTSTAT.

Then, using (21), we can conclude that

$$\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right)=0 \text{ only if } \left(\widehat{\boldsymbol{\rho}}(t),\left(h_i^B\left(\widehat{b}_i(t)\right)\right)_{i\in\mathcal{N}}\right)\in\mathcal{X}^{\pi} \quad (51)$$

Now, the above discussion along with (49), (50), and (51) allow us to conclude that for almost all t

$$\frac{dL\left(\widehat{\boldsymbol{\Theta}}(t)\right)}{dt}\leq\Delta\left(\widehat{\boldsymbol{\Theta}}(t)\right)$$
$$\text{where } \Delta\left(\widehat{\boldsymbol{\Theta}}\right)\leq 0\ \forall\ \boldsymbol{\Theta}\in\mathcal{H},\ \Delta(\boldsymbol{\Theta})<0\ \forall\ \boldsymbol{\Theta}\notin\mathcal{H}^*. \quad (52)$$

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

Proof of part (b) can be completed using (52) and an appropriate choice of a continuous function bounding $\Delta(\boldsymbol{\Theta})$ from above (obtained by replacing $\frac{1}{u_i(t)}$ with $\frac{1}{u_{\max}}$, where $u_{\max}$ is defined in (41), in the constituent functions of $\Delta(\boldsymbol{\Theta})$) so that the new function also satisfies (52) after replacement of $\Delta(\boldsymbol{\Theta})$ in (41). The details omitted for brevity and can be found in [21]. ■

References

[1] V. Joseph, G. de Veciana, Nova: Qoe-driven optimization of dash-based video delivery in networks, in: IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, 2014, pp. 82–90.

[2] C. Yim, A.C. Bovik, Evaluation of temporal variation of video quality in packet loss networks, Signal Process., Image Commun. (2011).

[3] T. Kim, M. Ammar, Optimal quality adaptation for scalable encoded video, IEEE J. Sel. Areas Commun. (2005).

[4] P. Seeling, M. Reisslein, B. Kulapala, Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial, IEEE Commun. Surv. Tutor. (2004).

[5] R. Mok, E. Chan, R. Chang, Measuring the quality of experience of HTTP video streaming, in: IFIP/IEEE International Symposium on Integrated Network Management, 2011.

[6] G. Ribezzo, G. Samela, V. Palmisano, L. De Cicco, S. Mascolo, A dash video streaming system for immersive contents, in: Proceedings of the 9th ACM Multimedia Systems Conference, MMSys '18, New York, NY, USA, Association for Computing Machinery, 2018, pp. 525–528.

[7] S.-C. Yen, C.-L. Fan, C.-H. Hsu, Streaming 360° videos to head-mounted virtual reality using dash over quic transport protocol, in: Proceedings of the 24th ACM Workshop on Packet Video, PV '19, New York, NY, USA, Association for Computing Machinery, 2019, pp. 7–12.

[8] M. Hosseini, View-aware tile-based adaptations in 360 virtual reality video streaming, in: 2017 IEEE Virtual Reality, VR, 2017, pp. 423–424.

[9] F. Fu, M. van der Schaar, A systematic framework for dynamically optimizing multi-user wireless video transmission, IEEE J. Sel. Areas Commun. (2010).

[10] J. Huang, Z. Li, M. Chiang, A. Katsaggelos, Joint source adaptation and resource allocation for multi-user wireless video streaming, IEEE Trans. Circuits Syst. Video Technol. (2008).

[11] H. Hu, X. Zhu, Y. Wang, R. Pan, J. Zhu, F. Bonomi, Qoe-based multi-stream scalable video adaptation over wireless networks with proxy, in: ICC, 2012.

[12] D. Bethanabhotla, G. Caire, M.J. Neely, Joint transmission scheduling and congestion control for adaptive video streaming in small-cell networks, 2013, arXiv:1304.8083 [Cs.NI].

[13] V. Joseph, G. de Veciana, Jointly optimizing multi-user rate adaptation for video transport over wireless systems: Mean-fairness-variability tradeoffs, in: IEEE INFOCOM, 2012.
[14] V. Joseph, S. Borst, M. Reiman, Optimal rate allocation for adaptive wireless video streaming in networks with user dynamics, in: IEEE INFOCOM, 2014.
[15] B. Blaszczyszyn, M. Jovanovic, M. Karray, Quality of real-time streaming in wireless cellular networks- Stochastic modeling and analysis, 2013, arXiv:1304.5034 [Cs.NI].
[16] S. Khan, Y. Peng, E. Steinbach, M. Sgroi, W. Kellerer, Application-driven cross-layer optimization for video streaming over wireless networks, IEEE Commun. Mag. (2006).
[17] S. Shakkottai, R. Srikant, Network optimization and control, Found. Trends Netw. (2007).
[18] V. Joseph, G. de Veciana, A. Arapostathis, Resource allocation: Realizing mean-variability-fairness tradeoffs, IEEE Trans. Automat. Control 60 (2015) 19–33.
[19] H. Kushner, G.G. Yin, Stochastic Approximation and Recursive Algorithms and Applications, Springer, 2003.
[20] A.L. Stolyar, Greedy primal–dual algorithm for dynamic resource allocation in complex networks, Queueing Syst. (2006).
[21] V. Joseph, Mean-Variability-Fairness Tradeoffs in Resource Allocation with Applications to Video Delivery (Ph.D. dissertation), ECE Department, UT Austin, 2013, Aug. http://users.ece.utexas.edu/gustavo/phd-theses/Joseph.Vinay2013.pdf.
[22] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2009.
[23] A.V. Fiacco, Y. Ishizuka, Sensitivity and stability analysis for nonlinear programming, Ann. Oper. Res. (1990).
[24] A. Fiacco, J. Kyparisis, Convexity and concavity properties of the optimal value function in parametric nonlinear programming, J. Optim. Theory Appl. (1986).
[25] J.F. Bonnans, A. Shapiro, Optimization problems with perturbations, a guided tour, SIAM Rev. (1996).
[26] A.L. Stolyar, On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation, Oper. Res. (2005).