Jointly Optimizing Multi-user Rate Adaptation for Video Transport over Wireless Systems: Mean-Fairness-Variability Tradeoffs

Vinay Joseph and Gustavo de Veciana
Department of Electrical and Computer Engineering, The University of Texas at Austin

***Abstract*—User perceived video quality depends on a variety of only partially understood factors, e.g., the application domain, content, compression, transport mechanism, and most importantly psycho-visual systems determining the ultimate Quality of Experience (QoE) of users. This paper centers on two key observations in addressing the problem of joint rate adaptation for video streams sharing a congested resource. First, we note that a user viewing a given video will experience temporal variations in the dependence of perceived video quality to the compression rate. Intuitively this is due to the possibly changing nature of the content, e.g., from an action to a slower scene. Thus, in allocating rates to users sharing a congested resource, in particular a wireless system where additional temporal variability in users' capacity may be high, content dependent tradeoffs can be realized to deliver a better overall average perceived video quality. Second, we note that such adaptation of users' rates, may result in temporal variations in video quality which combined with perceptual hysteresis effects will degrade users' QoE. We develop an asymptotically optimal online algorithm, requiring minimal statistical information, for optimizing users' QoE by realizing tradeoffs across mean, variance and fairness. Simulations show that our approach achieves significant gains in viewers' QoE.**

The novelty of this work lies not only in tackling the fundamental problem of achieving fair allocations of perceived video quality across a user population with time varying sensitivities and capacity, but, in addition, in integrating the deleterious impact that variations in perceived quality has on their QoE.

I. Introduction

There has been tremendous growth in the number of users viewing videos on mobile devices in the past decade. Current trends (see [5]) suggest that mobile video traffic will more than double each year between 2010-15, with two-thirds of mobile data traffic being video by 2015. It is unlikely that wireless infrastructure, e.g., base stations, access points, capacity etc., can keep up with such growth, and hence finding ways to make the most of available resources to deliver the best possible 'Quality of Experience' (QoE) to viewers is among the important networking problems today.

For a user viewing a video stream, the QoE associated with the session has a strong positive correlation to several metrics that increase with the *average* Perceived Video Quality (PVQ) across the sequence of scenes comprising the video – our convention will be that PVQ is a local measure associated with a particular scene or a short period of time. PVQ, in turn, depends on a variety of factors, including compression, network transport, content, human perceptual system, etc. A key observation, which is at center stage in this paper, is that the overall QoE also depends on *temporal variations* of PVQ across scenes, see e.g., [26], [13], [19]. Indeed [26] even points out that variations in PVQ can result in a QoE that is worse than that of a constant quality video with lower *average* PVQ.

There are several factors that can result in variations in PVQ for wireless users. We will focus on two prominent ones. The first, is the time varying nature of the wireless channel capacity due to fast fading (on faster time scales, e.g., ms) and slow fading due to shadowing, dynamic interference, mobility, and changing loads (on slower time scales, e.g. secs). Indeed 4G broadband systems promise to further increase extent and dynamic range of such variability [16]. The second, is the time varying nature of video scenes PVQ dependence on the (average) compression rate as well as other factors. Perhaps the key contributor to such change is the video content itself, e.g, the PVQ at a fixed compression rate might be smaller for an action scene (where there is a lot of changing visual content) than for a slower scene (where things stay the same).

Capturing the rich space of factors that impact PVQ for a sequence of scenes is challenging and perhaps impractical. In this paper we will abstract this as a (possibly) time varying sequence of (increasing) functions mapping the source compression rate to the PVQ for the current scene. Similar content-dependent functions were used in [10], [12] to map parameters like compression rate, compression scheme parameters, physical layer parameters etc to a measure of PVQ. Here we will abstract dependence of this large number of parameters, assuming source coder has already been optimized to deliver the best possible PVQ for a given compression rate or vice versa. This is possible, due to the availability of efficient, easy to use Video Quality Assessment algorithms (see [21] for a survey) such as the ones based on SSIM based indices (see [24], [25]) that can be used to efficiently evaluate how humans are likely to perceive video compressed in different ways and at different rates. For stored video these functions might be obtained and optimized offline, e.g., several compressed or layered (e.g., based on Scalable Video Coding [18]) versions of a video might be optimized to deliver the best relationships between PVQ and compressed rates. For video streaming of

This research was supported in part by Intel and Cisco under the VAWN program, and by the NSF under Grant CNS-0917067. We thank Zheng Lu for useful discussions and for help with simulations, and Anush Moorthy, Chen Chao and Kalpana Seshadrinathan for useful discussions.

live events, live broadcast of TV channels etc, experience with the content, or online use of computationally efficient Quality Assessment algorithms might be used to optimize PVQ-rate tradeoffs.

When multiple users share a 'congested' wireless resource and are subject to slow (e.g., secs) wireless capacity and PVQ-rate tradeoff variability, *joint source rate adaptation* will be key to determining users' QoE. As mentioned earlier such a scheme should attempt to fairly maximize users' (mean) PVQ while controlling its temporal variability. This is the focus of our paper. Note that fairness here is an additional system-level requirement to avoid excessively compromising the QoE of users that are either seeing poor channels, e.g., at the cell edge, and/or are currently resource (rate) intensive to achieve a reasonable PVQ.

A. Related Work

Several works have considered how to best deliver video to multiple users in a wireless setting, see e.g., [12], [23], [10], [27], [8]. In [12], the authors consider periodically and greedily maximizing the sum of the users' PVQ – peak signal-to-noise ratio is used as a crude measure for PVQ. The work of [23] focuses on a WLAN setting and proposes a scheme that greedily maximizes the minimum quality among users – perhaps an extreme fairness objective. This is done by determining the optimal encoding rate and physical layer parameters so as to minimize the sum of the distortion caused by source compression and the expected distortion resulting from packet loss during transmission. In [10] a distributed solution to the problem of minimizing the sum distortion is given– incurring increasing communication overhead. The works of [27] and [8] pose the problem in a dynamic programming framework which requires detailed knowledge of the statistics of the system. To remedy this, [8] proposes a learning algorithm to solve the associated Markov decision problem. In addition to complexity, and perhaps unreasonably detailed model assumptions, a major weakpoint of the above works is that none accounts for the impact of the variability of users' PVQ on their QoE.

There are several works, for e.g., [6], that propose schemes to reduce variability in coding rates to reduce the variability in PVQ. But, these approaches ignore the (time varying nature of the) dependence of PVQ on rates, and hence are suboptimal. The problem of reducing the variability in quality for SVC coded video over the Internet was considered in [13]. However, they restrict their attention to a single video stream, and they only focus on the reduction of switching rate of quality which is a crude metric for variability of PVQ.

B. Our Approach and Contributions

In this paper, we develop an approach for joint rate adaptation to maximize users' QoE by *fairly* increasing users' PVQ over time while controlling its variability. We first consider the optimal solution of an *offline* convex optimization problem (i.e., one where all time-varying quantities are assumed *known*) which essentially optimizes a time average of the sum of concave functions of PVQ of the users, minus its variance subject to the capacity constraints imposed by a wireless system. This is, in part, motivated by the QoE metric proposed in [26]. The main contribution of this paper is an *asymptotically optimal online* algorithm, AVQ, *requiring almost no statistical information* about the system, i.e., with sufficient time, the reward achieved by AVQ is arbitrarily close to that of the optimal offline solution. This is established under a general setting where temporal variations are stationary ergodic, but the algorithm could also be enabled to adapt to changing statistics. We then modify AVQ to obtain a *practical low complexity* algorithm PAVQ, recognizing the discrete nature of available compressed video content, e.g., SVC layers. Through simulation we show that PAVQ provides significant performance gains over a collection of known QoE metrics vs several baseline joint rate adaptation schemes.

C. Organization of the paper

In Section II, we discuss the system model. In Section III, we obtain and study the optimal offline finite horizon policy. In Section IV, we present AVQ and prove its optimality. In Section V, we present PAVQ. In Section VI, we evaluate the performance of PAVQ through simulations, and compare it against some other schemes. We conclude the paper in Section VII. The sketch of a proof of an intermediate result is given in Appendix B.

II. System model

We consider a wireless system serving video to a fixed set of users $\mathcal{N}$ where $|\mathcal{N}| = N$. We consider a slotted system where slots are indexed by $t \in \{0, 1, 2...\}$. We envisage the duration of a slot being longer than that of a Group of Pictures (roughly about a second) so that the evaluation of PVQ-rate tradeoffs for a slot make sense.

Throughout the paper, we distinguish between random variables (and random functions) and their realizations by using upper case letters for the former and lower case for the latter. We use bold letters to denote vectors, e.g., $\mathbf{a} = (a_i : i \in \mathcal{N})$. We let $(\mathbf{a})_{1:T}$ denote the finite length sequence $(\mathbf{a}(t) : 1 \leq t \leq T)$.

Let $C_i(t)$ be a random variable denoting the average rate at which user $i \in \mathcal{N}$ can be served in slot t. We assume that $\mathbf{C}(t) = (C_i(t) : i \in \mathcal{N})$ is known at the beginning of slot t. Knowledge of $\mathbf{C}(t)$ can be based on averaging short term estimates based on channel quality information or the achieved transmission rates over the wireless channel. These estimates are intended to capture slowly changing aspects of the channel like shadowing, interference, mobility etc. Further, we assume that there exist constants $c_{\min} > 0$ and $c_{\max} < \infty$, such that $C_i(t) \in [c_{\min}, c_{\max}]$ for all slots t and $i \in \mathcal{N}$ over all sample paths. Consider any sample path. Then, any rate adaptation scheme, while choosing the rates, $\mathbf{r}(t)$, at which users are served in slot t, should satisfy

$$r_i(t) \geq 0, \ \forall i \in \mathcal{N}, \text{ and } \sum_{i \in \mathcal{N}} \frac{r_i(t)}{c_i(t)} \leq 1, \ \forall \ t. \qquad (1)$$

Hence $r_i(t)/c_i(t)$ corresponds to the fraction of time the wireless system would have to serve user i in slot t to deliver $r_i(t)$ to user i. Thus the above constraint ensures that, in slot t the wireless system could deliver $\mathbf{r}(t)$ to the respective users.

For user i, let $F^Q_{t,i}(.)$ denote the (random) function, and $f^Q_{t,i}(.)$, its realization, that maps the user's source compression rate $r_i(t)$ to the user's PVQ $q_i(t)$ in slot t, i.e., $q_i(t) = f^Q_{t,i}(r_i(t))$. This function depends on the video content being sent to user i in slot t. We assume that these functions are picked from a finite (but potentially large) set $\mathcal{F}_{qual}$ where each function is a twice-differentiable increasing concave function that maps a video compression rate to a finite number representing the PVQ at that rate. As discussed in the introduction we envisage several scenarios (stored video and even live streaming) where PVQ-rate tradeoff functions can be known a priori or estimated each slot. Note since $r_i(t) \le c_i(t) \le c_{\max}$ for each $i \in \mathcal{N}$ and t, and $\mathcal{F}_{qual}$ is finite, there exists $q_{\max}$ such that $q_i(t) \le q_{\max}$ for all $i \in \mathcal{N}$ and in each slot t. Similarly we assume that there exists some $q_{\min} \ge 0$ such that for any realization of the channel $\mathbf{c}$, there exists $\mathbf{r}$ satisfying $\sum_{i\in\mathcal{N}} r_i/c_i < 1$ and $\min_{i\in\mathcal{N}} f^Q_{t,i}(r_i) \ge q_{\min}$. Since $c_{\min} > 0$, this condition can be satisfied by choosing arbitrarily small $q_{\min}$. We let $\mathcal{Q} = [q_{\min}, q_{\max}]$.

We assume a centralized coordinator, possibly the basestation, or colocated video optimizing server (see e.g., [16]) that has access to $\mathbf{S}(t) = \left(\mathbf{C}(t), \mathbf{F}^Q_t\right) \in \mathcal{S}$, where $\mathcal{S}$ denotes the set of values $\mathbf{S}(t)$ can take. The coordinator's role is to choose source rates (e.g., in practice this may translate to dropping SVC layers) so as to meet the current wireless capacity constraints. When this is the case, we assume that the associated PVQs are realized. Thus implicitly we assume negligible packet losses as would be achieved by a modern broadband wireless system through physical layer adaptation of modulation and coding and hybrid ARQ, with negligible delays, e.g., 200 ms (see [11]) which are deemed acceptable relative to video playback buffers for the applications mentioned in Section I.

III. Optimal Variance-Sensitive Offline Policy

In this section, we consider an offline formulation for optimal joint rate adaptation roughly maximizing video QoE subject to the wireless capacity constraints over a finite time horizon. In the offline setting we assume $\left(\mathbf{c}, \mathbf{f^Q}\right)_{1:T}$, i.e., the realization of the process $(\mathbf{S})_{1:T}$, is known. We consider maximizing the following objective function

$$\phi_T\left((\mathbf{q})_{1:T}\right) = \sum_{t=1}^{T}\sum_{i\in\mathcal{N}} U\left(q_i(t)\right) - \sum_{i\in\mathcal{N}}\left(\frac{\lambda_i T}{2}\right)\text{Var}\left((q_i)_{1:T}\right),$$

where

$$\text{Var}\left((q_i)_{1:T}\right) = \frac{1}{T}\sum_{t=1}^{T}\left(q_i(t) - \frac{1}{T}\sum_{\tau=1}^{T} q_i(\tau)\right)^2.$$

The first term in ϕ_T increases in users' PVQ $(\mathbf{q})_{1:T}$ which in turn depends on the allocated rates $(\mathbf{r})_{1:T}$, while the second term penalizes variability of the PVQ. The objective function is closely related to the QoE metric proposed in [26] except that we have substituted the standard deviation of PVQ by the variance. Additionally, $U : \mathcal{Q} \to \mathbb{R}_+$ is a strictly increasing concave function which serves to enforce some fairness in the allocation of quality amongst users in each slot. For instance, we can choose U from the following class of strictly concave increasing functions parametrized by $\alpha \in (0, \infty)$ ([15])

$$U_\alpha(q) = \begin{cases} \log q & \text{if } \alpha = 1, \\ (1-\alpha)^{-1} q^{1-\alpha} & \text{otherwise.} \end{cases} \quad (2)$$

If we set $U = U_\alpha$, a larger α corresponds to a more fair allocation of quality. In the sequel, we also assume that U is twice differentiable. Here, $\boldsymbol{\lambda} = (\lambda_i : i \in \mathcal{N})$ are positive parameters that are chosen to roughly reflect the importance given to the reduction of temporal variability of PVQ. Let $\lambda_{\max} = \max_{i\in\mathcal{N}} \lambda_i$.

We consider the optimization problem OPT(T) given below:

$$\max_{(\mathbf{q})_{1:T}} \phi_T\left((\mathbf{q})_{1:T}\right) \quad (3)$$

$$\text{subject to} \quad \sum_{i\in\mathcal{N}} \frac{f^R_{t,i}(q_i(t))}{c_i(t)} \le 1 \quad \forall\, t \in \{1, ..., T\}, \quad (4)$$

$$q_i(t) \ge q_{\min} \forall\, t \in \{1, ..., T\}, \forall\, i \in \mathcal{N}, \quad (5)$$

where $f^R_{t,i}(.)$ is the inverse function of $f^Q_{t,i}(.)$. The constraints (4) and (5) ensure that (1) is not violated. We choose $q_{\min}$ as described in Section II.

Lemma 1 below asserts that OPT(T) is a strictly convex optimization problem satisfying Slater's condition (Section 5.2.3, [4]), and thus has a unique solution. Proving this is straightforward once one establishes that Var$((\mathbf{q})_{1:T})$ is a convex function and is 'almost' a strictly convex function of $(\mathbf{q})_{1:T}$.

Lemma 1. *OPT(T) is a convex optimization problem satisfying Slater's condition with a unique solution.*

Proof: The inverse of an increasing concave function is a convex function, and hence (4) is a convex constraint. Now consider the objective function, $\phi_T()$. The first term in $\phi_T()$ is clearly a concave function. Now, consider the second term. For two quality vectors $\left(\mathbf{q^1}\right)_{1:T}$ and $\left(\mathbf{q^2}\right)_{1:T}$, any $i \in \mathcal{N}$, $\alpha \in (0, 1)$ and $\bar{\alpha} = 1 - \alpha$, we have that

$$\text{Var}\left(\alpha\left(q^1_i\right)_{1:T} + \bar{\alpha}\left(q^2_i\right)_{1:T}\right)$$

$$
\begin{aligned}
&= \mathrm{Var}\left(\left(\alpha q_i^1 + \bar{\alpha} q_i^2\right)_{1:T}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\left(\alpha q_i^1(t) + \bar{\alpha} q_i^2(t)\right)\right. \\
&\quad \left. -\frac{1}{T}\sum_{\tau=1}^{T}\left(\alpha q_i^1(\tau) + \bar{\alpha} q_i^2(\tau)\right)\right)^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\alpha\left(q_i^1(t) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^1(\tau)\right)\right. \\
&\quad \left. +\bar{\alpha}\left(q_i^2(t) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^2(\tau)\right)\right)^2 \\
&\leq \frac{1}{T}\sum_{t=1}^{T}\left(\alpha\left(q_i^1(t) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^1(\tau)\right)^2\right. \\
&\quad \left. +\bar{\alpha}\left(q_i^2(t) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^2(\tau)\right)^2\right) \\
&= \alpha \mathrm{Var}\left(\left(q_i^1\right)_{1:T}\right) + \bar{\alpha}\mathrm{Var}\left(\left(q_i^2\right)_{1:T}\right).
\end{aligned}
$$

The inequality in the penultimate line follows from the convexity of $f(x) = x^2$. We conclude that Var() is a convex function. Further, since $f(x) = x^2$ is a strictly convex function, the inequality is a strict one unless

$$q_i^1(t) = q_i^2(t) + \frac{1}{T}\sum_{\tau=1}^{T} q_i^1(\tau) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^2(\tau) \ \forall\ 1 \leq t \leq T.$$

Thus, for the inequality not to be a strict one, we require that $\mathrm{Var}\left(\left(q_i^1\right)_{1:T}\right) = \mathrm{Var}\left(\left(q_i^2\right)_{1:T}\right)$.

Using the above arguments and concavity of U, we conclude that OPT(T) is a convex optimization problem.

To establish uniqueness of the optimal solution, let $\left(\mathbf{q^1}\right)_{1:T}$ and $\left(\mathbf{q^2}\right)_{1:T}$ be two optimal solutions to OPT(T). Then, from the concavity of the objective, $\left(\alpha\left(q_i^1\right)_{1:T} + \bar{\alpha}\left(q_i^2\right)_{1:T}\right)$ is also an optimal solution for any $\alpha \in (0, 1)$ and $\bar{\alpha} = 1 - \alpha$. Due to concavity of U and convexity of Var(), this is only possible if for each $i \in \mathcal{N}$ and $1 \leq t \leq T$, $U\left(\alpha q_i^1(t) + \bar{\alpha} q_i^2(t)\right) = \alpha U\left(q_i^1(t)\right) + \bar{\alpha} U\left(q_i^2(t)\right)$, and $\mathrm{Var}\left(\alpha\left(q_i^1\right)_{1:T} + \bar{\alpha}\left(q_i^2\right)_{1:T}\right) = \alpha\mathrm{Var}\left(\left(q_i^1\right)_{1:T}\right) + \bar{\alpha}\mathrm{Var}\left(\left(q_i^2\right)_{1:T}\right)$. From above discussion, $\mathrm{Var}\left(\alpha\left(q_i^1\right)_{1:T} + \bar{\alpha}\left(q_i^2\right)_{1:T}\right) = \alpha\mathrm{Var}\left(\left(q_i^1\right)_{1:T}\right) + \bar{\alpha}\mathrm{Var}\left(\left(q_i^2\right)_{1:T}\right)$ for each $i \in \mathcal{N}$ is only possible if $\mathrm{Var}\left(\left(q_i^1\right)_{1:T}\right) = \mathrm{Var}\left(\left(q_i^2\right)_{1:T}\right)$ for each $i \in \mathcal{N}$, and $q_i^1(t) = q_i^2(t) + \frac{1}{T}\sum_{\tau=1}^{T} q_i^1(\tau) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^2(\tau)$ for each $i \in \mathcal{N}$ and $1 \leq t \leq T$. Since for each $i \in \mathcal{N}$, $\mathrm{Var}\left(\left(q_i^1\right)_{1:T}\right) = \mathrm{Var}\left(\left(q_i^2\right)_{1:T}\right)$, due to optimality of $\left(\mathbf{q^1}\right)_{1:T}$ and $\left(\mathbf{q^2}\right)_{1:T}$, we have that

$$
\begin{aligned}
&\sum_{1\leq t\leq T}\sum_{i\in\mathcal{N}} U\left(q_i^2(t)\right) = \sum_{1\leq t\leq T}\sum_{i\in\mathcal{N}} U\left(q_i^1(t)\right) \\
&= \sum_{1\leq t\leq T}\sum_{i\in\mathcal{N}} U\left(q_i^2(t) + \frac{1}{T}\sum_{\tau=1}^{T} q_i^1(\tau) - \frac{1}{T}\sum_{\tau=1}^{T} q_i^2(\tau)\right).
\end{aligned}
$$

Since U is a strictly increasing function, the above equation implies that

$$\frac{1}{T}\sum_{\tau=1}^{T} q_i^1(\tau) = \frac{1}{T}\sum_{\tau=1}^{T} q_i^2(\tau),$$

and thus,

$$q_i^1(t) = q_i^2(t)\ \forall\ 1 \leq t \leq T,\ \forall\ i \in \mathcal{N}.$$

From the above discussion, we can conclude that OPT(T) has a unique solution.

Slater's condition is satisfied due to the careful choice of $q_{\min}$. ■

We let $\left(\mathbf{q}^T\right)_{1:T}$ denote the optimal solution to OPT(T), and note that once optimal quality allocations are given, the associated optimal rate allocations, denoted by $\left(\mathbf{r}^T\right)_{1:T}$, are given by $r_i^T(t) = f_{t,i}^R(q_i^T(t))$.

Since OPT(T) is a convex optimization problem satisfying Slater's condition (Lemma 1), Karush-Kuhn-Tucker (KKT) conditions ([4]) are necessary and sufficient for optimality. Specifically, we have that $\left(\mathbf{q^T}\right)_{1:T}$ is an optimal solution to OPT(T) if and only if it is feasible, and there exist non-negative constants $\left(\mu^T\right)_{1:T}$ and $\left(\gamma_i^T : i \in \mathcal{N}\right)_{1:T}$ such that for all $i \in \mathcal{N}$ and $t \in \{1, ..., T\}$

$$
\begin{aligned}
U'\left(q_i^T(t)\right) - \lambda_i\left(q_i^T(t) - \overline{q}_i^T\right) & \\
-\mu^T(t)\frac{\left(f_{t,i}^R\right)'\left(q_i^T(t)\right)}{c_i(t)} + \gamma_i^T(t) &= 0, \quad (6) \\
\mu^T(t)\left(\sum_{i\in\mathcal{N}}\frac{f_{t,i}^R(q_i^T(t))}{c_i(t)} - 1\right) &= 0, \quad (7) \\
\gamma_i^T(t)\left(q_i^T(t) - q_{\min}\right) &= 0, \quad (8)
\end{aligned}
$$

and

$$\overline{q}_i^T = \frac{1}{T}\sum_{t=1}^{T} q_i^T(t).$$

Here we have used the fact that for any $i \in \mathcal{N}$ and $\tau' \in \{1, ..., T\}$

$$\frac{\partial}{\partial q_i(\tau')}\left(T\mathrm{Var}\left((q_i)_{1:T}\right)\right) = 2\left(q_i(\tau') - \frac{1}{T}\sum_{\tau=1}^{T} q_i(\tau)\right).$$

From (6), for all $i \in \mathcal{N}$ and $t \in \{1, 2, ..., T\}$

$$\mu^T(t) = \frac{\lambda_i\left(\overline{q}_i^T - q_i^T(t)\right) + U'\left(q_i^T(t)\right) + \gamma_i^T(t)}{\left(\left(f_{t,i}^R\right)'\left(q_i^T(t)\right)/c_i(t)\right)}. \quad (9)$$

Thus, for optimality, in each slot t one must ensure that the RHS of the above equation is equal for all users $i \in \mathcal{N}$. A key observation here is that, for $t_1 \neq t_2$, $\mathbf{q}^T(t_1)$ and $\mathbf{q}^T(t_2)$ are only related through $\overline{\mathbf{q}}^T$. So, if a genie revealed $\overline{\mathbf{q}}^T$, the optimal quality allocation $\mathbf{q}^T(t)$ for each slot t, can be determined by solving an optimization depending only on the information about the current slot, i.e., $\mathbf{S}(t)$. The online algorithm proposed in next section exploits this idea.

IV. Adaptive Variance aware Quality allocation

In this section, we present our algorithm, AVQ, and establish its asymptotic optimality.
AVQ consists of three steps, AVQ.0-AVQ.2, given next:

AVQ.0: Initialize: Let $\widehat{q}_i^*(0) \in \mathcal{Q}$ for each $i \in \mathcal{N}$.

In each slot t, carry out the following steps:
AVQ.1: The quality allocation in slot t is the optimal solution to the optimization problem OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t)$) given below:

$$\max_{\mathbf{q}} \sum_{i\in\mathcal{N}} U\left(q_i\right) - \sum_{i\in\mathcal{N}} \left(\frac{\lambda_i}{2}\right)\left(q_i - \widehat{q}_i^*(t-1)\right)^2 \quad (10)$$

$$\text{subject to} \quad \sum_{i\in\mathcal{N}} \frac{f_{t,i}^R(q_i)}{c_i(t)} \leq 1, \quad (11)$$

$$q_i \geq q_{\min} \ \forall\ i \in \mathcal{N}. \quad (12)$$

Let $\mathbf{q}^*(\widehat{\mathbf{q}}, \mathbf{s})$ denote the solution to OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$).
AVQ.2: In slot t, update $\widehat{q}_i$ as follows: for all $i \in \mathcal{N}$,

$$\begin{aligned} \widehat{q}_i^*(t) &= \widehat{q}_i^*(t-1) \quad (13)\\ &+ \frac{\lambda_i}{t+\lambda_{\max}}\left(q_i^*(\widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t)) - \widehat{q}_i^*(t-1)\right). \end{aligned}$$

We see that AVQ.1 is quite intuitive by comparing (10) to (3). In place of user i's variance in (3), $\left(q_i - \widehat{q}_i^*(t-1)\right)^2$ appears in (10) in each slot. In the sequel, we show that $\left(\widehat{q}_i^*\right)_t$ converges to the mean of user i's quality allocation over time.

Throughout the paper, for brevity, we use $\mathbf{q}^*(t)$ instead of $\mathbf{q}^*(\widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t))$ when the dependence of $\mathbf{q}^*(t)$ on $(\widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t))$ is clear from context.

For any t, OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t)$) is a convex optimization problem satisfying Slater's condition. So, the optimal solution satisfies KKT conditions and thus, there exist non-negative constants $\mu^*(t)$ and $(\gamma_i^*(t) : i \in \mathcal{N})$ such that for all $i \in \mathcal{N}$

$$\begin{aligned} U^{'}\left(q_i^*(t)\right) - \lambda_i\left(q_i^*(t) - \widehat{q}_i^*(t-1)\right) & \\ + \gamma_i^*(t) - \mu^*(t)\frac{\left(f_{t,i}^R\right)^{'}\left(q_i^*(t)\right)}{c_i(t)} &= 0, \quad (14)\\ \mu^*(t)\left(\sum_{i\in\mathcal{N}} \frac{f_{t,i}^R(q_i^*(t))}{c_i(t)} - 1\right) &= 0, \quad (15)\\ \gamma_i^*(t)\left(q_i^*(t) - q_{\min}\right) &= 0. \quad (16) \end{aligned}$$

The rest of this section is devoted to proving the asymptotic optimality of AVQ. We use intermediate results, Lemmas 2-5, to prove the main result given in Theorem 1.

A. Convergence Analysis

In this subsection, we show under fairly weak assumptions on the process $\left(\mathbf{S}(t)\right)_t$, $\left(\widehat{\mathbf{q}}^*(t)\right)_t$ converges almost surely. This is a key intermediate result in proving the main optimality result discussed in the next subsection.

We begin with the next result in which we establish continuity and differentiability properties of $\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ and $h\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ respectively, as functions of $\widehat{\mathbf{q}}$, where

$$h(\widehat{\mathbf{q}}, \mathbf{s}) = \sum_{i\in\mathcal{N}} U\left(q_i^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right)\right) - \sum_{i\in\mathcal{N}} \frac{\lambda_i}{2}\left(q_i^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right) - \widehat{q}_i\right)^2.$$

Note that $\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ is the optimizer of OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$). So, $h(\widehat{\mathbf{q}}, \mathbf{s})$ is the value of the optimized objective function. See Appendix A for a proof of the result which mainly relies on some fundamental results on perturbation analysis of optimization problems from [7] and [3], and Bounded Convergence Theorem (see [9]).

Lemma 2. *For any* $\mathbf{s} \in \mathcal{S}$,
(a) $\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ *is a continuous function of* $\widehat{\mathbf{q}}$*;*
(b) For each $i \in \mathcal{N}$, $\left(\nabla_{\widehat{\mathbf{q}}} h\left(\widehat{\mathbf{q}}, \mathbf{s}\right)\right)_i = -\lambda_i\left(q_i^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right) - \widehat{q}_i\right)$.
Let $\left(\mathbf{S}(t)\right)_t$ *be a stationary ergodic process. Then*
(c) $E\left[\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right]$ *is a continuous function of* $\widehat{\mathbf{q}}$*;*
(d) For each $i \in \mathcal{N}$,
$\left(\nabla_{\widehat{\mathbf{q}}} E\left[h\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right]\right)_i = -\lambda_i\left(E\left[q_i^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right] - \widehat{q}_i\right)$.

In the sequel, we require the following assumption to hold for the functions in $\mathcal{F}_{qual}$ and U .

Assumption 1. *At least one of the following holds:*
A1*: There exists positive constant* $\delta_{U''}$ *such that* $U^{''}(q) \leq -\delta_{U''}\ \forall\ q \in \mathcal{Q}$.
A2*: There exist positive constants* $\delta_{f''}$, $\delta_{f'}$ *and* $\delta_{U'}$ *such that for any inverse function* f *of* $f^Q \in \mathcal{F}_{qual}$,

$$f^{''}(q) \geq \delta_{f''},\ f^{'}(q) \leq \delta_{f'} \text{ and } U^{'}(q) \geq \delta_{U'} \quad \forall\ q \in \mathcal{Q}.$$

Later in the section, we show that $\left(\widehat{\mathbf{q}}^*(t)\right)_t$ converges to the fixed point of (17). The uniqueness of its fixed point is discussed in the next result, and the proof of the result is sketched in Appendix B.

Lemma 3. *Let* $\left(\mathbf{S}(t)\right)_t$ *be a stationary ergodic process. Under Assumption 1, the following fixed point equation has a unique solution*

$$E\left[\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right] = \widehat{\mathbf{q}}. \quad (17)$$

Next we show that, under a fairly weak assumption on $\left(\mathbf{S}(t)\right)_t$ given below, $\left(\widehat{\mathbf{q}}^*(t)\right)_t$ converges.

Assumption 2. *At least one of the following holds:*
S1*:* $\left(\mathbf{S}(t)\right)_t$ *is an i.i.d. process.*
S2*:* $\left(\mathbf{S}(t)\right)_t$ *is a stationary ergodic process taking values in a finite* $\mathcal{S}$.

The next result says that $\left(\widehat{\mathbf{q}}^*(t)\right)_t$ converges if Assumptions 1 and 2 hold. A proof of the result is given in Appendix C. The proof mainly proceeds by viewing (13) as a stochastic approximation update equation, and using results from [14] that give sufficient conditions for convergence of a stochastic approximation scheme.

Lemma 4. *Suppose Assumptions 1 and 2 hold. Then, if $\widehat{\mathbf{q}}^*(0) \in \mathcal{Q}^N$, the sequence $(\widehat{\mathbf{q}}^*(t))_t$ generated by AVQ converges almost surely to $\widehat{\mathbf{q}}$, the unique fixed point of*

$$E\left[\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right] = \widehat{\mathbf{q}}.$$

To prove the asymptotic optimality of AVQ, we need the following intermediate result which essentially relies on the convergence of $(\widehat{\mathbf{q}}^*(t))_t$.

Lemma 5. *Suppose Assumptions 1 and 2 hold. Then, for each $i \in \mathcal{N}$, the following limits converge almost surely*

$$\lim_{T\to\infty}\left(\frac{1}{T}\sum_{t=1}^{T} q_i^*(t) - \widehat{q}_i^*(T)\right) = 0.$$

Proof: Consider any realization of $(\mathbf{S}(t))_t$. For any $i \in \mathcal{N}$, it follows from (13) that

$$\sum_{t=1}^{T}\left(\frac{t+\lambda_{\max}}{\lambda_i}\right)\left(\widehat{q}_i^*(t) - \widehat{q}_i^*(t-1)\right) = \sum_{t=1}^{T}\left(q_i^*(t) - \widehat{q}_i^*(t-1)\right).$$

Noting the telescopic sums in the above expression, we can simplify it to obtain

$$\frac{T\widehat{q}_i^*(T) - \sum_{t=1}^{T}\widehat{q}_i^*(t-1)}{\lambda_i} - \frac{\lambda_{\max}}{\lambda_i}\left(\widehat{q}_i^*(T) - \widehat{q}_i^*(1)\right) = \sum_{t=1}^{T}\left(q_i^*(t) - \widehat{q}_i^*(T) + \widehat{q}_i^*(T) - \widehat{q}_i^*(t-1)\right).$$

From Lemma 4, we know that if Assumptions 1 and 2 hold, $(\widehat{\mathbf{q}}^*(t))_t$ converges almost surely. Hence, limits of the terms in the LHS of the equation below exist, and

$$\lim_{T\to\infty}\frac{\widehat{q}_i^*(T) - \frac{1}{T}\sum_{t=1}^{T}\widehat{q}_i^*(t-1)}{\lambda_i} - \lim_{T\to\infty}\left(\widehat{q}_i^*(T) - \frac{1}{T}\sum_{t=1}^{T}\widehat{q}_i^*(t-1)\right) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left(q_i^*(\widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t)) - \widehat{q}_i^*(T)\right).$$

Since $(\widehat{\mathbf{q}}^*(t))_t$ converges, the above equation implies that for almost all sample paths

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left(q_i^*(t) - \widehat{q}_i^*(T)\right) = 0.$$

■

B. Asymptotic Optimality of AVQ

The next result establishes the asymptotic optimality of AVQ, i.e., if we run AVQ for long enough period, the time average of the difference in performance of AVQ and the optimal finite horizon policy becomes negligible. *This is a strong result as it is comparing AVQ, an online algorithm, against the optimal offline scheme which has access to $(\mathbf{s})_{1:T}$, i.e., channel and $\left(\mathbf{f}^R\right)_{1:T}$, ahead of time.*

Theorem 1. *Suppose Assumptions 1 and 2 hold. Then, for almost all sample paths*

$$\lim_{T\to\infty}\frac{1}{T}\left(\phi_T\left((\mathbf{q}^*)_{1:T}\right) - \phi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right)\right) = 0.$$

Proof: Consider any realization of $(\mathbf{s})_{1:T}$. Let $(\mu^*)_{1:T}$ and $(\gamma_i^* : i \in \mathcal{N})_{1:T}$ be the sequences of non negative real numbers satisfying (14), (15) and (16) for the realization. Hence, from the non-negativity of these numbers, and feasibility of $\left(\mathbf{q}^T\right)_{1:T}$, we have

$$\phi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \quad \leq \quad \varphi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right).$$

where

$$\begin{aligned}\varphi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \quad = \quad & \sum_{t=1}^{T}\sum_{i\in\mathcal{N}} U\left(q_i^T(t)\right) \\ & - \sum_{i\in\mathcal{N}}\left(\frac{\lambda_i T}{2}\right)\text{Var}\left(\left(q_i^T\right)_{1:T}\right) \\ & - \sum_{t=1}^{T}\mu^*(t)\left(\sum_{i\in\mathcal{N}}\frac{f_{t,i}^R(q_i^T(t))}{c_i(t)} - 1\right) \\ & + \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\gamma_i^*(t)\left(q_i^T(t) - q_{\min}\right).\end{aligned}$$

Since φ_T is a differentiable concave function, we have (see [4])

$$\begin{aligned}\varphi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \quad \leq \quad & \varphi_T\left((\mathbf{q}^*)_{1:T}\right) \\ + \quad & \nabla\varphi_T\left((\mathbf{q}^*)_{1:T}\right) \bullet \left(\left(\mathbf{q}^T\right)_{1:T} - (\mathbf{q}^*)_{1:T}\right),\end{aligned}$$

where '•' denotes the dot product. Hence, we have

$$\begin{aligned}
&\phi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \le \varphi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \\
&\le \sum_{t=1}^{T}\sum_{i\in\mathcal{N}} U\left(q_i^*(t)\right) - \sum_{i\in\mathcal{N}}\left(\frac{\lambda_i T}{2}\right)\mathrm{Var}\left(\left(q_i^*\right)_{1:T}\right) \\
&\quad - \sum_{t=1}^{T}\mu^*(t)\left(\sum_{i\in\mathcal{N}}\frac{f_{t,i}^R(q_i^*(t))}{c_i(t)} - 1\right) \\
&\quad + \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\gamma_i^*(t)\left(q_i^*(t) - q_{\min}\right) \\
&\quad + \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\left(q_i^T(t) - q_i^*(t)\right) \\
&\quad \left(U^{'}\left(q_i^*(t)\right) + \frac{\lambda_i}{T}\sum_{\tau=1}^{T} q_i^*(\tau) - \lambda_i q_i^*(t)\right. \\
&\quad \left. -\mu^*(t)\frac{\left(f_{t,i}^R\right)^{'}\left(q_i^*(t)\right)}{c_i(t)} + \gamma_i^*(t)\right).
\end{aligned}$$

Now, since $(\mu^*)_{1:T}$ and $(\gamma_i^* : i\in\mathcal{N})_{1:T}$ satisfy (14), (15) and (16), we have

$$\begin{aligned}
\phi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \;\le\; & \sum_{t=1}^{T}\sum_{i\in\mathcal{N}} U\left(q_i^*(t)\right) \\
- \; & \sum_{i\in\mathcal{N}}\left(\frac{\lambda_i T}{2}\right)\mathrm{Var}\left(\left(q_i^*\right)_{1:T}\right) \\
+ \; & \sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\lambda_i\left(q_i^T(t) - q_i^*(t)\right) \\
& \left(\frac{1}{T}\sum_{\tau=1}^{T} q_i^*(\tau) - \widehat{q}_i^*(t-1)\right). \quad (18)
\end{aligned}$$

Consider the following expression appearing in the last term in the above line:

$$\begin{aligned}
&\left(\frac{1}{T}\sum_{\tau=1}^{T} q_i^*(\tau)\right) - \widehat{q}_i^*(t-1) \\
&= \widehat{q}_i^*(T) - \widehat{q}_i^*(t-1) + \left(\frac{1}{T}\sum_{\tau=1}^{T} q_i^*(\tau)\right) - \widehat{q}_i^*(T).
\end{aligned}$$

From Lemma 5, and since $\left(\widehat{\mathbf{q}}^*(t)\right)_t$ converges under Assumptions 1 and 2 (Lemma 4), we know that we can make the above term as small as required by choosing a large enough t for almost all sample paths. Also, $\left|q_i^T(t) - q_i^*(t)\right| \le q_{\max}$. Thus, for almost all sample paths

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\lambda_i\left(q_i^T(t) - q_i^*(t)\right) \\
&\qquad\left(\left(\frac{1}{T}\sum_{\tau=1}^{T} q_i^*(\tau)\right) - \widehat{q}_i^*(t-1)\right) = 0.
\end{aligned}$$

Hence, taking limits in (18),

$$\lim_{T\to\infty}\frac{1}{T}\left(\phi_T\left(\left(\mathbf{q}^*\right)_{1:T}\right) - \phi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right)\right) \ge 0.$$

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

$$\phi_T\left(\left(\mathbf{q}^T\right)_{1:T}\right) \ge \phi_T\left(\left(\mathbf{q}^*\right)_{1:T}\right).$$

From the above two inequalities, the result follows. ■

V. PAVQ: Practical Algorithm Motivated by AVQ

In current practical settings, content providers typically set a *discrete* set of video source coding rates or quality levels to be achieved, e.g., for the associated SVC layers. Note that if quality levels are fixed, then the associated rates might vary with the content, and vice versa, over slots, effectively giving time varying PVQ-rate tradeoffs. This limits the choice of quality/rate allocations. So AVQ, which assumes PVQ as (time varying) 'nice' (continuous, differentiable etc.) functions of rate, cannot be directly applied in such settings. In this section, we present PAVQ, a joint rate adaptation algorithm that does quality allocations emulating AVQ's approach, which takes into account these practical considerations.

Let $\mathcal{R}_{t,i}$ denote the set of rates at which user i's video can be encoded in slot t. In slot t, let $r_{t,i}(.)$ denote the strictly increasing function that maps rate indices in $\{1, 2, ..., |\mathcal{R}_{t,i}|\}$ to coding rates in $\mathcal{R}_{t,i}$ for user $i \in \mathcal{N}$. The algorithm PAVQ is described below.

PAVQ.0: Initialize: For all $i\in\mathcal{N}$, let $\widehat{q}_i^P(0) \in \mathcal{Q}$.

In each slot t, carry out the following steps:
PAVQ.1: The quality allocation in slot t to user i, $q_i^P(t)$, is given by $f_{t,i}^Q\left(r_{t,i}\left(\rho_i\right)\right)$ where ρ_i is the output of the following algorithm:

Initialize: $\rho_i = 1$ for all $i\in\mathcal{N}$; $\mathcal{I} = \{1, 2..., N\}$;
while $\mathcal{I} \neq \{\}$
  For each $i\in\mathcal{I}$, evaluate

$$\mu_i^P = \frac{\lambda_i\left(\widehat{q}_i^p(t-1) - f_{t,i}^Q\left(r_{t,i}\left(\rho_i\right)\right)\right) + U^{'}\left(f_{t,i}^Q\left(r_{t,i}\left(\rho_i\right)\right)\right)}{\left(\frac{1}{c_i(t)}\right)\frac{r_{t,i}(\rho_i+1) - r_{t,i}(\rho_i)}{f_{t,i}^Q(r_{t,i}(\rho_i+1)) - f_{t,i}^Q(r_{t,i}(\rho_i))}};$$

  if $\max_{i\in\mathcal{I}}\mu_i^P < 0$
    $\mathcal{I} = \{\}$;
  else
    $i^* = \mathrm{argmax}_{i\in\mathcal{I}}\left(\mu_i^P\right)$;
    $\rho_{i^*} = \rho_{i^*} + 1$;
    if $\rho_{i^*} = |\mathcal{R}_{t,i^*}|$
      $\mathcal{I} = \mathcal{I} \setminus \{i^*\}$;
    end
    if $\sum_{i\in\mathcal{N}}\frac{r_{t,i}(\rho_i)}{c_i(t)} > 1$
      $\rho_{i^*} = \rho_{i^*} - 1$;
      $\mathcal{I} = \mathcal{I} \setminus \{i^*\}$;
    end
  end
end

PAVQ.2: In slot t, we update $\widehat{q}_i^P$ as follows: for each $i\in\mathcal{N}$

$$\widehat{q}_i^p(t) = \widehat{q}_i^P(t-1) + \frac{\lambda_i}{t+\lambda_{\max}}\left(q_i^P(t) - \widehat{q}_i^p(t-1)\right). \quad (19)$$

Instead of finding the optimal quality allocations using AVQ.2, PAVQ finds good quality allocations $\left(\mathbf{q}^P(t)\right)_t$ using PAVQ.2. The reason for using μ_i^P in PAVQ.2 is clear from (14) (using (14), we can obtain an expression similar to (9) for $\mu^*(t)$) by comparing the μ_i^P to $\mu^*(t)$. Also, the number of loops associated with the 'while' in PAVQ.1 is at most $\sum_{i\in\mathcal{N}}|\mathcal{R}_{t,i}|$ in slot t. So the complexity of PAVQ is roughly proportional to the product of the number of users times the number of rates per user.

We now introduce two additional joint rate allocation schemes, PMUR and PMUQ, to be used for comparison purposes in Section VI. The first, PMUR, corresponds to a strategy that allocates rates to users with time varying wireless capacities in a fair manner without considering PVQ-rate dependencies. It performs quality allocation in each slot t by running PAVQ.1 with one change: we replace Step 2 of PAVQ.1 with

02: For each $i \in \mathcal{I}$, evaluate

$$\mu_i^P = \frac{U'\left(r_{t,i}\left(\rho_i\right)\right)}{\left(\frac{1}{c_i(t)}\right)\frac{r_{t,i}(\rho_i+1)-r_{t,i}(\rho_i)}{f^Q_{t,i}(r_{t,i}(\rho_i+1))-f^Q_{t,i}(r_{t,i}(\rho_i))}};$$

Following a reasoning similar to that for the choice of μ_i^P in PAVQ.1, we can verify that PMUR is roughly carrying out the following optimization in each slot t (for some small $r_{\min}$):

$$\max_{\mathbf{r}\in\mathbb{R}^N}\left\{\sum_{i\in\mathcal{N}}U\left(r_i\right):\sum_{i\in\mathcal{N}}\frac{r_i}{c_i(t)}\le 1,\ r_i\ge r_{\min}\ \forall i\right\}.$$

The second, PMUQ, corresponds to a strategy that allocates PVQ fairly amongst users with time varying PVQ-rate tradeoffs and wireless capacities. However, it ignores temporal variability in PVQ. It carries out quality allocation in each slot t by running PAVQ.1 with one change: we replace Step 2 of PAVQ.1 with

02: For each $i \in \mathcal{I}$, evaluate

$$\mu_i^P = \frac{U'\left(f^Q_{t,i}\left(r_{t,i}\left(\rho_i\right)\right)\right)}{\left(\frac{1}{c_i(t)}\right)\frac{r_{t,i}(\rho_i+1)-r_{t,i}(\rho_i)}{f^Q_{t,i}(r_{t,i}(\rho_i+1))-f^Q_{t,i}(r_{t,i}(\rho_i))}};$$

Following a reasoning similar to that for the choice of μ_i^P in PAVQ.1, we can verify that PMUQ is roughly carrying out the following optimization in each slot t:

$$\max_{\mathbf{q}\in\mathbb{R}^N}\left\{\sum_{i\in\mathcal{N}}U\left(q_i\right):\sum_{i\in\mathcal{N}}\frac{f^R_{t,i}(q_i)}{c_i(t)}\le 1,\ q_i\ge q_{\min}\ \forall i\right\}.$$

VI. Simulations

In this section, we use simulations to compare the performance of PAVQ, to that of PMUR and PMUQ in different scenarios with $U=\log$ and $T=10^5$ slots.

We consider four scenarios:

(a) I.I.D. channels and homogeneous users;
(b) I.I.D. channels and heterogeneous users;
(c) Markov channels and homogeneous users;
(d) Markov channels and heterogeneous users.

Under scenarios (a) and (b), users see I.I.D. channels: $(C_i(t))_t$ for each user $i\in\mathcal{N}$ is obtained independently in each slot t from a distribution which is representative of capacities seen by a randomly placed user with single antenna equalizer in an HSDPA system with 50% load (and thus associated interference) from its neighbors. Under scenarios (c) and (d), users see Markov channels: we use Markov Chain Monte Carlo method to generate $(C_i(t))_t$ in such a way that the values in consecutive slots are positively correlated, and yet the stationary distribution matches the above distribution. Furthermore in scenarios (b) and (d), we consider heterogeneous users: a third of the users see $0.5\,(C)_{1:T}$, a third of the users see $(C)_{1:T}$ and remaining third see $1.5\,(C)_{1:T}$, where $(C)_{1:T}$ is the channel seen by a typical user under (a) over T slots.

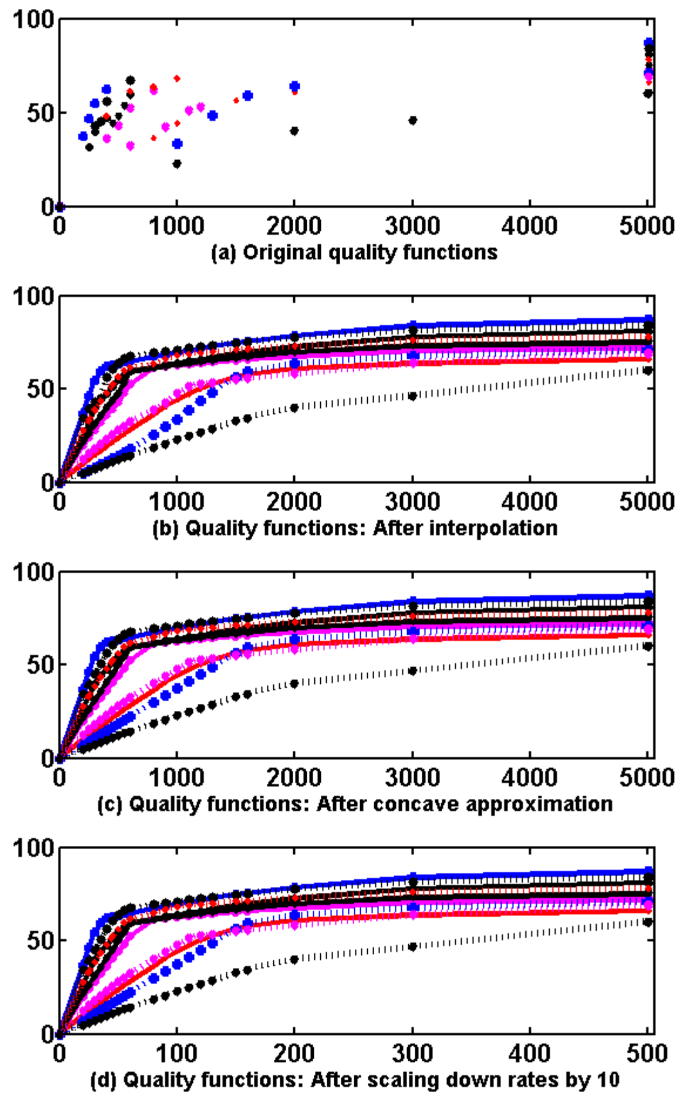

Fig. 1. Obtaining functions in $\mathcal{F}_{qual}$

In all the scenarios, the PVQ-rate functions, $F^Q_{t,i}$ are chosen independently and uniformly at random from $\mathcal{F}_{qual}$ in each slot t for each user $i \in \mathcal{N}$. Fig. 1 (a) shows estimates of perceived quality at various rates that were obtained using data from LIVE Video Quality Database ([22]), and represent PVQ-rate tradeoffs for 8 sec segments of diverse compressed videos. Then, using cubic interpolation (using MATLAB), we found approximate values of perceived quality at certain rates of interest (See Fig. 1 (b)). Then, we obtained a concave approximation of the functions by taking these function (See Fig. 1 (c)). The values in Fig. 1 (a) were obtained for videos watched on devices with big screens over the wireline networks, and hence the rates involved are fairly large. We scale down the rates by 10 to get reasonable approximations for applications used by mobile devices that typically do not need such high rates (See Fig. 1 (d)).

We compare the performance of PMUR and PMUQ against that of PAVQ$_{10^{-3}}$, which is PAVQ with $\lambda_i = 10^{-3}$ for all $i \in \mathcal{N}$, and PAVQ$_1$, which is PAVQ with $\lambda_i = 1$ for all $i \in \mathcal{N}$. For each algorithm π, we consider following metrics:

$$\begin{aligned}
\overline{Q} &= \frac{1}{NT}\sum_{i\in\mathcal{N}}\sum_{t=1}^{T} q_i^{\pi}(t),\\
\text{Var} &= \frac{1}{N}\sum_{i\in\mathcal{N}} \text{Var}\left((q_i^{\pi})_{1:T}\right),\\
\overline{\wedge Q} &= \frac{T_F}{NT}\sum_{i\in\mathcal{N}}\sum_{t=1}^{T/T_F} \wedge q_i^{\pi}(t),\\
M_K &= \frac{1}{N}\sum_{i\in\mathcal{N}} M_{Ki},\\
\overline{|\Delta Q|} &= \frac{1}{N(T-1)}\sum_{i\in\mathcal{N}}\sum_{t=2}^{T} \left|q_i^{\pi}(t) - q_i^{\pi}(t-1)\right|,\\
\text{CV} &= \frac{1}{N}\sum_{i\in\mathcal{N}}\left(\frac{\sqrt{\text{Var}\left((q_i^{\pi})_{1:T}\right)}}{\text{Mean}\left((q_i^{\pi})_{1:T}\right)}\right),
\end{aligned}$$

where $\wedge q_i^{\pi}(t) = \min_{\tau}\left(q_i^{\pi}(\tau) : \tau \in [(t-1)T_F + 1, tT_F]\right)$, $T_F = 10$ and M_{Ki} is the QoE metric proposed in [20] (setting parameters $\tau = 2$, $K = 50$, $\alpha = 0.8$).

Instead of comparing our schemes using a single QoE metric, we will use all the above metrics – there is no 'perfect' QoE metric in the literature and the search for one is an ongoing work. Note $\overline{Q}$, CV and Var are averages across users of the mean, coefficient of variation and variance of the PVQ seen over time. Coefficient of variation (CV) is the ratio of the standard deviation and mean, and has an impact on QoE (see for e.g., [19]). We use $\overline{\wedge Q}$ to roughly capture the worst perceived quality seen by the users over time which has a major impact on the QoE (as pointed out in [20]). $\overline{|\Delta Q|}$ captures the average change in perceived quality of the users over consecutive slots. Finally, note that using $\overline{Q}$ and Var, one can compute QoE metric proposed in [26] for scenario (a).

For almost all combinations of N and the scenarios, Tables (I)-(IV) show that PMUQ outperforms PMUR in all the metrics. This was expected as the former utilizes the knowledge of the dependence of quality on the coding rates.

We can also see that PAVQ$_1$ outperforms the algorithms as far the metrics $\overline{\wedge Q}$, M_K, $\overline{|\Delta Q|}$, CV and Var are concerned. For instance, we see that CV and Var of PAVQ$_1$ is significantly less than that of PMUQ. This was expected as our problem formulation (OPT(T)) explicitly accounts for the impact of PVQ variance. Though we only considered variance in our formulation, we see that the PAVQ$_1$ has significantly lesser $\overline{|\Delta Q|}$ than PMUQ. For instance, in scenario (d) (see Table IV), the percentage reduction of $\overline{|\Delta Q|}$ ranges from 42%-64%. Also, the percentage gains over PMUQ in the QoE metric M_K ranges from 20%-48% under scenario (a) (see Table I).

As expected, PMUQ has better $\overline{Q}$ when compared to PAVQ$_1$ in all almost all settings. But, one can control the emphasis on reducing variability by reducing $(\lambda_i)_{i\in\mathcal{N}}$. Indeed PAVQ$_{10^{-3}}$ has better $\overline{Q}$ though it comes at the cost of other metrics.

In the simulations, we observed that $(\widehat{\mathbf{q}})_t$ converges fairly quickly. For instance, consider Fig. 2 which depicts $(\widehat{q}_i)_{1:500}$ for $i \in \{1, 11, 21\}$ under Setting (b) for PAVQ$_1$ with $N = 30$. The value to which $(\widehat{q}_i)_t$ converges is different for different $i \in \{1, 11, 21\}$ due to heterogeneity of the users in Setting (b), e.g., user corresponding to $i = 21$ sees better channels on average.

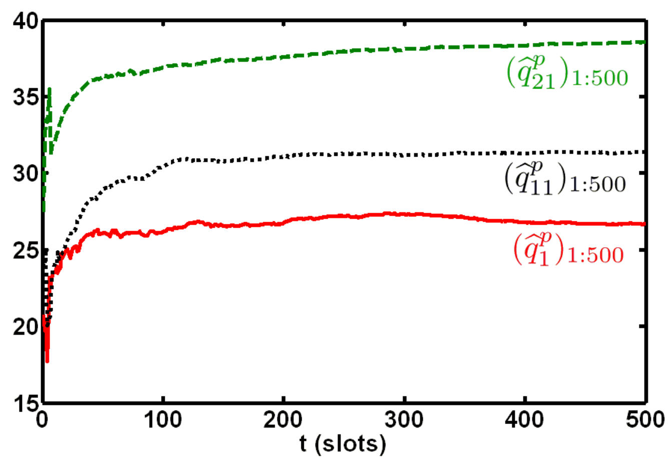

Fig. 2. Convergence of $(\widehat{q}_i)_{1:500}$ for $i \in \{1, 11, 21\}$

VII. Conclusions

We proposed first centralized joint rate adaptation algorithm which is sensitive to variability in the perceived video quality. It outperforms aggressive baseline solutions across multiple QoE metrics. The approach has the merit of allowing the engineer to realize tradeoffs in fairness across users, mean PVQ and variability in the PVQ. Our approach captures the key sources of variability in such systems and essential requirements of such systems, and yet is surprisingly spare in terms of requiring a priori statistical information.

TABLE I
SCENARIO (A): IID CHANNELS, HOMOGENEOUS USERS

ALG, N	$\overline{Q}$	$\overline{\wedge Q}$	M_K	$\overline{\lvert\Delta Q\rvert}$	CV, Var
PAVQ$_1$, 15	49.3	**41.2**	**32.9**	**3.8**	**0.10, 23.1**
PAVQ$_{10^{-3}}$, 15	58.2	40.2	29.5	9.0	0.16, 85
PMUR, 15	56.8	23.7	15.6	20.1	0.32, 339.4
PMUQ, 15	**58.8**	38.2	27.5	11.2	0.18, 119.5
PAVQ$_1$, 30	34.3	**24.2**	**16.9**	**4.8**	**0.15, 27.0**
PAVQ$_{10^{-3}}$, 30	43.0	19.4	12.5	14.4	0.30, 172
PMUR, 30	44.2	13.2	8.9	23.4	0.46, 419.2
PMUQ, 30	**45.3**	17.1	11.4	18.3	0.36, 272.8
PAVQ$_1$, 45	26.8	**15.3**	**10.0**	**7.0**	**0.25, 43.2**
PAVQ$_{10^{-3}}$, 45	32.4	11.9	8.0	14.2	0.39, 156
PMUR, 45	35.1	9.5	6.8	22.5	0.56, 384.32
PMUQ, 45	**35.3**	10.6	7.4	20.0	0.49, 301.8

TABLE II
SCENARIO (B): IID CHANNELS, HETEROGENEOUS USERS

ALG, N	$\overline{Q}$	$\overline{\wedge Q}$	M_K	$\overline{\lvert\Delta Q\rvert}$	CV, Var
PAVQ$_1$, 15	46.3	19.6	**29.5**	**3.9**	**0.11, 24.4**
PAVQ$_{10^{-3}}$, 15	55.8	**19.9**	25.9	10.3	0.19, 109
PMUR, 15	54.5	14.7	15.1	20.3	0.35, 342.6
PMUQ, 15	**56.5**	18.4	24.2	12.4	0.21, 146.4
PAVQ$_1$, 30	32.6	**18.2**	**15.4**	**5.5**	**0.19, 33.8**
PAVQ$_{10^{-3}}$, 30	40.3	16.1	11.9	13.8	0.32, 159
PMUR, 30	41.5	12.0	8.7	22.2	0.49, 383.5
PMUQ, 30	**42.3**	14.0	10.5	18.3	0.40, 269.7
PAVQ$_1$, 45	25.1	**12.3**	**8.1**	**8.5**	**0.32, 61.2**
PAVQ$_{10^{-3}}$, 45	29.3	10.3	7.1	13.7	0.42, 146
PMUR, 45	**31.4**	8.6	6.3	20.0	0.56, 317.5
PMUQ, 45	**31.4**	9.3	6.7	18.4	0.52, 260.2

TABLE III
SCENARIO (C): MARKOV CHANNELS, HOMOGENEOUS USERS

ALG, N	$\overline{Q}$	$\overline{\wedge Q}$	M_K	$\overline{\lvert\Delta Q\rvert}$	CV, Var
PAVQ$_1$, 15	50.7	22.0	**41.5**	**3.4**	**0.09, 21.0**
PAVQ$_{10^{-3}}$, 15	57.7	**22.4**	38.1	7.8	0.16, 88
PMUR, 15	56.2	18.1	28.6	15.2	0.33, 342.5
PMUQ, 15	**59.8**	21.8	39.3	9.1	0.17, 109.1
PAVQ$_1$, 30	36.4	**22.8**	**24.9**	**4.9**	**0.17, 39.7**
PAVQ$_{10^{-3}}$, 30	43.5	22.4	22.8	10.6	0.30, 170
PMUR, 30	43.7	17.3	16.8	17.0	0.47, 409.4
PMUQ, 30	**45.1**	20.2	20.4	13.7	0.36, 266.7
PAVQ$_1$, 45	27.5	**19.0**	**16.6**	**6.4**	**0.25, 47.0**
PAVQ$_{10^{-3}}$, 45	32.8	17.3	14.4	10.9	0.39, 163
PMUR, 45	**35.5**	13.7	11.5	17.0	0.55, 378.7
PMUQ, 45	35.2	15.4	12.7	15.0	0.49, 294.8

TABLE IV
SCENARIO (D): MARKOV CHANNELS, HETEROGENEOUS USERS

ALG, N	$\overline{Q}$	$\overline{\wedge Q}$	M_K	$\overline{\lvert\Delta Q\rvert}$	CV, Var
PAVQ$_1$, 15	48.2	24.2	**38.8**	**3.4**	**0.10, 23.1**
PAVQ$_{10^{-3}}$, 15	56.3	**25.6**	37.2	8.2	0.18, 98
PMUR, 15	55.8	21.2	29.0	15.0	0.33, 324.3
PMUQ, 15	**57.0**	23.6	35.5	10.1	0.2 142.5
PAVQ$_1$, 30	33.1	21.8	**22.6**	**5.2**	**0.20, 38.0**
PAVQ$_{10^{-3}}$, 30	40.5	**22.2**	20.6	10.6	0.32, 158
PMUR, 30	41.7	17.6	16.2	16.6	0.48, 375.6
PMUQ, 30	**43.4**	20.5	19.7	13.5	0.38, 258.7
PAVQ$_1$, 45	25.2	14.9	**12.8**	**8.2**	**0.34, 69.8**
PAVQ$_{10^{-3}}$, 45	30.4	**15.3**	**12.8**	11.0	0.41, 150
PMUR, 45	**31.7**	11.9	10.0	15.8	0.56, 310.9
PMUQ, 45	31.6	13.2	10.9	14.3	0.51, 258.6

In future work, we expect to build upon this framework towards fully characterizing potential gains and tradeoffs that joint rate adaptation can provide in such systems, and explore the possibility of further simplifying our approach allowing a distributed implementation.

REFERENCES

[1] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc., 1989.

[2] F. Bonnans, J. F. Bonnans, and A. D. Ioffe. Quadratic growth and stability in convex programming problems with multiple solutions, 1995.

[3] J. F. Bonnans and A. Shapiro. Optimization problems with perturbations, a guided tour. *SIAM REVIEW*, 40:228–264, 1996.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2009.

[5] CISCO. Cisco visual networking index: Global mobile data traffic forecast update, 2010-2015. February 2011.

[6] P. de Cuetos and K. W. Ross. Adaptive rate control for streaming stored fine-grained scalable video. In *Proceedings of NOSSDAV*, 2002.

[7] A. V. Fiacco and Y. Ishizuka. Sensitivity and stability analysis for nonlinear programming. *Annals of Operations Research*, 1990.

[8] F. Fu and M. van der Schaar. A systematic framework for dynamically optimizing multi-user wireless video transmission. *IEEE J.Sel. A. Commun.*, 28, April 2010.

[9] B. Hajek. An exploration of random processes for engineers. July 2011.

[10] J. Huang, Z. Li, M. Chiang, and A. Katsaggelos. Joint source adaptation and resource allocation for multi-user wireless video streaming. *IEEE Transactions on Circuits and Systems for Video Technology*, May 2008.

[11] J. C. Ikuno, M. Wrulich, and M. Rupp. Performance and modeling of LTE H-ARQ. In *Proceedings of WSA 2009*. EURASIP, 2009.

[12] S. Khan, Y. Peng, E. Steinbach, M. Sgroi, and W. Kellerer. Application-driven cross-layer optimization for video streaming over wireless networks. *Communications Magazine, IEEE*, Jan. 2006.

[13] T. Kim and M. Ammar. Optimal quality adaptation for scalable encoded video. *IEEE Journal on Selected Areas in Communications*, Deb. 2005.

[14] H. J. Kushner and G. G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[15] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Netw.*, October 2000.

[16] ORTIVA. Optimizing mobile video. July 2010.

[17] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., second edition.

[18] H. Schwarz, D. Marpe, and T. Wiegand. Overview of the scalable video coding extension of the H.264/AVC standard. *IEEE Transactions on Circuits and Systems for Video Technology*, Sept. 2007.

[19] P. Seeling, M. Reisslein, and B. Kulapala. Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial. *Communications Surveys Tutorials, IEEE*, 2004.

[20] K. Seshadrinathan and A. Bovik. Temporal hysteresis model of time varying subjective video quality. *ICASSP*, May 2011.

[21] K. Seshadrinathan and A. C. Bovik. Automatic prediction of perceptual quality of multimedia signals–a survey. *Multim. Tools Appl.*, Jan. 2011.

[22] K. Seshadrinathan, R. Soundararajan, A. Bovik, and L. Cormack. Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, June 2010.

[23] G.-M. Su, Z. Han, M. Wu, and K. J. R. Liu. Joint uplink and downlink optimization for real-time multiuser video streaming over wlans. *Circuits and Systems for Video Technology, IEEE Transactions on*, August 2007.

[24] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, april 2004.

[25] Z. Wang, L. Lu, and A. C. Bovik. Video quality assessment based on structural distortion measurement. *Sig. Proc. Image Comm.*, Feb. 2004.
[26] C. Yim and A. C. Bovik. Evaluation of temporal variation of video quality in packet loss networks. *Image Commun.*, January 2011.
[27] X. Zhu and B. Girod. Analysis of multi-user congestion control for video streaming over wireless networks. *Multimedia and Expo, IEEE International Conference on*, 2006.

Appendix A
Proof of Lemma 2

Proof: We will use Theorem 2.2 from [7] to establish part (a) of this lemma by verifying that OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$) satisfies the conditions in of the theorem. Theorem 2.2 (i) is satisfied since the objective function of OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$) is a continuous function of $\mathbf{q}$ for any $\widehat{\mathbf{q}}$. Theorem 2.2 (ii) is satisfied since the feasible region of OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$) is a closed set, and independent of $\widehat{\mathbf{q}}$. Theorem 2.2 (iii) is satisfied since OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$) is always feasible, and because the square term in the objective function ensures strict concavity which in turn ensures that the optimal solution is unique. Theorem 2.2 (iv) is satisfied since $\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ is uniformly compact near $\widehat{\mathbf{q}}$. This is due to the fact that $\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ for any $\widehat{\mathbf{q}}$ lies in the bounded set $\mathcal{Q}^N$. Thus, all conditions of Theorem 2.2 hold, and part (a) follows.

Part (b) of the lemma follows from Theorem 4.1 in [3] and the remark following the theorem.

To verify part (c), consider any sequence $(\widehat{\mathbf{q}}_m)_m$ such that $\lim_{m\to\infty} \widehat{\mathbf{q}}_m = \widehat{\mathbf{q}}$, we have

$$\begin{aligned}\lim_{m\to\infty} E\left[q_i^*\left(\widehat{\mathbf{q}}_m, \mathbf{S}(t)\right)\right] &= E\left[\lim_{m\to\infty} q_i^*\left(\widehat{\mathbf{q}}_m, \mathbf{S}(t)\right)\right] \\ &= E\left[q_i^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right],\end{aligned}$$

which follows from Bounded Convergence Theorem since $q_i^* \leq q_{\text{max}}$, and part (a). This establishes part (c).

To prove (d), we show that for each $i \in \mathcal{N}$,

$$\frac{\partial}{\partial \widehat{q}_i} E\left[h\left(\widehat{\boldsymbol{q}}, \mathbf{S}(t)\right)\right] = E\left[\frac{\partial}{\partial \widehat{q}_i} h\left(\widehat{\boldsymbol{q}}, \mathbf{S}(t)\right)\right].$$

For any sequence $(\widehat{q}_{im})_m$ such that $\lim_{m\to\infty} \widehat{q}_{im} = 0$, using Mean Value Theorem and Lemma 2 (b), for $\mathbf{s} \in \mathcal{S}$, we have

$$\begin{aligned}&\left|\frac{h\left(\widehat{\mathbf{q}} + \mathbf{e}_i \widehat{q}_{im}, \mathbf{s}\right) - h\left(\widehat{\mathbf{q}}, \mathbf{s}\right)}{\widehat{q}_{im}}\right| \\ &\qquad = \lambda_i \left|q_i^*\left(\widehat{\mathbf{q}} + \mathbf{e}_i q_{0m}, \mathbf{s}\right) - \widehat{q}_i - q_{0m}\right| \\ &\qquad \leq \lambda_i q_{\text{max}},\end{aligned}$$

for $0 < q_{0m} < \widehat{q}_{im}$. Hence, from Bounded Convergence Theorem, we have

$$\begin{aligned}&\lim_{m\to\infty} E\left[\frac{h\left(\widehat{\mathbf{q}} + \mathbf{e}_i \widehat{q}_{im}, \mathbf{S}(t)\right) - h\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)}{\widehat{q}_{im}}\right] \\ &\quad = E\left[\lim_{m\to\infty}\left(\frac{h\left(\widehat{\mathbf{q}} + \mathbf{e}_i \widehat{q}_{im}, \mathbf{S}(t)\right) - h\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)}{\widehat{q}_{im}}\right)\right] \\ &\quad = -\lambda_i E\left[q_i^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right) - \widehat{q}_i\right].\end{aligned}$$

■

Appendix B
Proof of Lemma 3

Proof: We begin the proof by introducing some new notation. Given any $\mathbf{q} \in \mathcal{Q}^N$, let $\widetilde{\mathbf{q}}$ be defined as follows: $\widetilde{q}_i = \sqrt{\lambda_i} q_i$ for each $i \in \mathcal{N}$. Also define function $\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}}, \mathbf{s}\right)$ as follows: $\widetilde{q_i^*}\left(\widetilde{\widehat{\mathbf{q}}}, \mathbf{s}\right) = \sqrt{\lambda_i} q_i^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right) \ \forall\ i \in \mathcal{N}$. Essentially $\widetilde{\ }$ is an overloaded operator that operates on vectors and the functions $\left(q_i^* : i \in \mathcal{N}\right)$. Also, for any $\mathbf{q}^1, \mathbf{q}^2 \in \mathbb{R}^N$, let $d\left(\mathbf{q}^1, \mathbf{q}^2\right) = \sqrt{\sum_{i\in\mathcal{N}} \left(q_i^1 - q_i^2\right)^2}$.

To prove the result, we will show two intermediate results. The first is that the following fixed point equation has at least one solution

$$E\left[\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}}, \mathbf{S}(t)\right)\right] = \widetilde{\widehat{\mathbf{q}}}. \tag{20}$$

Secondly, we show that $E\left[\widetilde{\mathbf{q}^*}\left(., \mathbf{S}(t)\right)\right]$ is a pseudo-contraction ([1]) which implies the uniqueness of fixed point.

Indeed, $E\left[\widetilde{\mathbf{q}^*}\left(., \mathbf{S}(t)\right)\right]$ is a continuous function (See Lemma 2 (c)) mapping a convex compact subset of $\mathbb{R}^N$ to itself. So, the existence of a fixed point of (20) follows from Brouwer's fixed point theorem (see [17]).

To show that $E\left[\widetilde{\mathbf{q}^*}\left(., \mathbf{S}(t)\right)\right]$ is a pseudo-contraction, we prove the following claim: Assume that Assumption 1 holds, and let $\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}$ be a solution of (20). Then for any $\widehat{\mathbf{q}}^{\mathbf{2}} \in \mathcal{Q}^N$ such that $\widehat{\mathbf{q}}^{\mathbf{2}} \neq \widehat{\mathbf{q}}^{\mathbf{1}}$, we have

$$d\left(E\left[\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right)\right], E\left[\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right]\right) < d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right).$$

Since, $\mathcal{Q}^N$ is a compact set, the above result implies that $E\left[\widetilde{\mathbf{q}^*}\left(., \mathbf{S}(t)\right)\right]$ is a pseudo-contraction.

To prove the claim, we first show that $\widetilde{\mathbf{q}^*}\left(., \mathbf{s}\right)$ is Lipschitz continuous with a Lipschitz constant less than one. Note that for $\mathbf{s} = \left(\mathbf{c}, \mathbf{f}^R\right)$, $\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}}, \mathbf{s}\right)$ is the optimal solution for M-OPTAVQ($\boldsymbol{\lambda}, \widetilde{\widehat{\mathbf{q}}}, \mathbf{s}$). a modification of OPTAVQ($\boldsymbol{\lambda}, \widehat{\mathbf{q}}, \mathbf{s}$) (see (10)-(12)), given below:

$$\begin{aligned}&\max_{\widetilde{\mathbf{q}}} h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}}\right) \\ &\text{subject to} \quad \sum_{i\in\mathcal{N}} \frac{1}{c_i} f_i^R\left(\frac{\widetilde{q}_i}{\sqrt{\lambda_i}}\right) \leq 1, \\ &\qquad\qquad\quad \frac{\widetilde{q}_i}{\sqrt{\lambda_i}} \geq q_{\text{min}}\ \forall\ i \in \mathcal{N},\end{aligned} \tag{21}$$

where

$$h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}}\right) = \sum_{i\in\mathcal{N}} U\left(\frac{\widetilde{q}_i}{\sqrt{\lambda_i}}\right) - \sum_{i\in\mathcal{N}} \frac{\lambda_i}{2}\left(\frac{\widetilde{q}_i}{\sqrt{\lambda_i}} - \frac{\widetilde{\widehat{q}}_i}{\sqrt{\lambda_i}}\right)^2.$$

We use Proposition 6.1 from [3] to verify the Lipschitz continuity (and obtain the Lipschitz constant) of $\widetilde{\mathbf{q}^*}\left(., \mathbf{s}\right)$ in a neighborhood of $\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}$. We verify that the two conditions given in the proposition hold. The first condition requires that the function $\Delta h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right)$ given below is Lipschitz continuous in a neighborhood of $\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}$:

$$\begin{aligned}\Delta h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right) &= h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right) - h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}\right). \\ &= \frac{1}{2}\sum_{i\in\mathcal{N}}\left(\widetilde{\widehat{q}}_i^1 - \widetilde{\widehat{q}}_i^2\right)\left(2\widetilde{q}_i - \widetilde{\widehat{q}}_i^1 - \widetilde{\widehat{q}}_i^2\right).\end{aligned}$$

Then,

$$\begin{aligned}&\left|\Delta h_0\left(\widetilde{\mathbf{q}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right) - \Delta h_0\left(\widetilde{\mathbf{q}^{\mathbf{2}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right)\right| \\ &\qquad = \left|\sum_{i\in\mathcal{N}}\left(\widetilde{\widehat{q}}_i^1 - \widetilde{\widehat{q}}_i^2\right)\left(\widetilde{q_i^1} - \widetilde{q_i^2}\right)\right| \\ &\qquad \le d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right) d\left(\widetilde{\mathbf{q}^{\mathbf{1}}}, \widetilde{\mathbf{q}^{\mathbf{2}}}\right), \quad (22)\end{aligned}$$

where the last step follows from Cauchy-Schwarz inequality. Hence, the first condition given in Proposition 6.1, [3] holds.

Next, we show that the second condition referred to as second-order growth condition given in the proposition also holds. The condition requires that there exists a positive constant c such that

$$h_0\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right), \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}\right) - h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}\right) \ge c\left(d\left(\widetilde{\mathbf{q}}, \widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)\right)\right)^2,$$

for each feasible (for M-OPTAVQ($\boldsymbol{\lambda}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}$)) $\widetilde{\mathbf{q}}$ in a neighborhood of $\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)$ To check this condition, we apply Theorem 6.1 from [2] to M-OPTAVQ($\boldsymbol{\lambda}, \widetilde{\widehat{\mathbf{q}}}, \mathbf{s}$) with $\widetilde{\widehat{\mathbf{q}}} = \widetilde{\widehat{\mathbf{q}}}^1$. The theorem considers convex optimization problems and provides sufficient conditions for the second order growth condition (Theorem 6.1 (v)) to hold when Slater qualification hypothesis holds (see [2]). It is not hard to see that Slater qualification hypothesis holds due to the our choice of q_{min} (See Section II). Now, we verify that one of the sufficient conditions, Theorem 6.1 (vii), is satisfied. For this, let

$$\begin{aligned}&L\left(\widetilde{\mathbf{q}}, \mu, (\gamma_i : i\in\mathcal{N})\right) \\ &= h_0\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right), \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}\right) - h_0\left(\widetilde{\mathbf{q}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}\right) \\ &\quad + \mu\left(\sum_{i\in\mathcal{N}}\frac{1}{c_i}f_i^R\left(\frac{\widetilde{q}_i}{\sqrt{\lambda_i}}\right) - 1\right) \\ &\quad - \sum_{i\in\mathcal{N}}\gamma_i\left(\frac{\widetilde{q}_i}{\sqrt{\lambda_i}} - q_{\text{min}}\right).\end{aligned}$$

Then, for any $\mathbf{d}\in\mathbb{R}^N$, the function ψ in Theorem 6.1 (vii) is given by

$$\psi_{\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}(\mathbf{d}) = \mathbf{d}^{tr}\nabla^2_{\widetilde{\mathbf{q}}}L\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right), \mu^m, (\gamma_i^m : i\in\mathcal{N})\right)\mathbf{d}$$

where μ^m and $(\gamma_i^m : i\in\mathcal{N})$ are the optimal Lagrange multipliers, $\nabla^2_{\widetilde{\mathbf{q}}}$ denotes the Hessian taken with respect to $\widetilde{\mathbf{q}}$, and $\mathbf{d}^{tr}$ is the transpose of the vector $\mathbf{d}$. We can show that

$$\psi_{\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}(\mathbf{d}) = \sum_{i\in\mathcal{N}} d_i^2\left(1 - \frac{1}{\lambda_i}U''\left(\frac{\widetilde{q_i^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}{\sqrt{\lambda_i}}\right) + \frac{\mu^m}{\lambda_i}\frac{\left(f_{t,i}^R\right)''\left(\frac{\widetilde{q_i^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}{\sqrt{\lambda_i}}\right)}{c_i}\right)$$

for any $\mathbf{d}\in\mathbb{R}^N$. Also, since μ^m is an optimal Lagrange multiplier (satisfying KKT conditions), $\mu^m = 0$ if (21) is not active, and otherwise $\forall\ i\in\mathcal{N}$ satisfies

$$\mu^m \ge c_i\left(\frac{U'\left(\frac{\widetilde{q_i^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}{\sqrt{\lambda_i}}\right) - \sqrt{\lambda_i}\left(\widetilde{q_i^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right) - \widetilde{\widehat{q}}_i^1\right)}{\left(f_{t,i}^R\right)'\left(\frac{\widetilde{q_i^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}{\sqrt{\lambda_i}}\right)}\right). \quad (23)$$

Hence, we can conclude that for any $\mathbf{d}\in\mathbb{R}^N$,

$$\psi_{\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}(\mathbf{d}) \ge \sum_{i\in\mathcal{N}} d_i^2. \quad (24)$$

Further, if A1 in Assumption 1 holds, then

$$\psi_{\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}(\mathbf{d}) \ge \left(1 + \frac{\delta_{U''}}{\lambda\text{max}}\right)\sum_{i\in\mathcal{N}} d_i^2.$$

Thus, Theorem 6.1 (vii) is satisfied and hence from Theorem 6.1 of [2], Theorem 6.1 (v) holds, i.e., second order growth condition holds. Thus, both the conditions require for Proposition 6.1 of [3] are satisfied, and using the proposition, we can conclude that under A1, for any $\mathbf{s}\in\mathcal{S}$,

$$\begin{aligned}d\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right), \widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{s}\right)\right) &\le \left(1 + \frac{\delta_{U''}}{\lambda\text{max}}\right)^{-1} d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right) \\ &< d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right).\end{aligned}$$

Thus, under A1, we can conclude that

$$E\left[\left(d\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right), \widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right)\right)^2\right] < \left(d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right)\right)^2.$$

Now, suppose A2 holds. Pick any $j\in\mathcal{N}$. Since $\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}$ satisfies (20), $\widetilde{q_j^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right) \le \widetilde{\widehat{q}}_j^1$ with some probability $p_j > 0$. Since, U is a strictly increasing function, $\widetilde{q_j^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right) \le \widetilde{\widehat{q}}_j^1$ implies that (21) is active, and hence from (23), we have

$$\mu^m \ge c_j(t)\left(\frac{U'\left(\frac{\widetilde{q^*}_j\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}{\sqrt{\lambda_j}}\right)}{\left(f_{j,t}^R\right)'\left(\frac{\widetilde{q^*}_j\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right)}{\sqrt{\lambda_j}}\right)}\right) \ge c_{\text{min}}\frac{\delta_{U'}}{\delta_{f'}}.$$

Thus, with probability atleast p_j,

$$\psi_{\widetilde{\mathbf{q}}^*}(\mathbf{d}) \geq \sum_{i\in\mathcal{N}} d_i^2 \left(1 + \frac{c_{\min}\delta_{U'}}{\lambda_i \delta_{f'}} \frac{\left(f_{t,i}^R\right)'' \left(\frac{\widetilde{q^*}_i\left(\widetilde{\widehat{\mathbf{q}}}^{\mathbf{1}}, \mathbf{S}(t)\right)}{\sqrt{\lambda_i}}\right)}{C_i(t)}\right) \geq \left(1 + \frac{c_{\min}\delta_{U'}\delta_{f''}}{\lambda_{\max}\delta_{f'} c_{\max}}\right) \sum_{i\in\mathcal{N}} d_i^2.$$

Thus, with probability atleast p_j, Theorem 6.1 (vii) is satisfied and hence from Theorem 6.1 of [2], Theorem 6.1 (v) holds, i.e., second order growth condition holds. Thus, both the conditions given in Proposition 6.1 of [3] are satisfied, and using the proposition, we can conclude that under A2, with probability at least p_j,

$$d\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right), \widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right) < d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right).$$

We can also use (24) along with similar arguments to apply Proposition 6.1 of [3] to show that for any $\mathbf{s}$,

$$d\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{s}\right), \widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{s}\right)\right) \leq d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right).$$

Thus, under A2 also, we can conclude that

$$E\left[\left(d\left(\widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right), \widetilde{\mathbf{q}^*}\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right)\right)^2\right] < \left(d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right)\right)^2.$$

Thus, if A1 or A2 holds,

$$\begin{aligned} & d\left(E\left[\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right)\right], E\left[\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right]\right) \\ &= \sqrt{\sum_{i\in\mathcal{N}}\left(E\left[\widetilde{q}_i^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right) - \widetilde{q}_i^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right]\right)^2} \\ &\leq \sqrt{\sum_{i\in\mathcal{N}} E\left[\left(\widetilde{q}_i^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right) - \widetilde{q}_i^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right)^2\right]} \\ &= \sqrt{E\left[\left(d\left(\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}\right)\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right)\right)\right)^2\right]} \\ &< d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right). \end{aligned}$$

The third step follows from the Jensen's inequality. This proves the claim and hence, the second intermediate result.

Now, to show the main result, suppose (20) has more than one solution. Let $\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}$ and $\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}$ be two distinct solutions of (20). Then, from part above result, we have

$$\begin{aligned} d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right) &= d\left(E\left[\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \mathbf{S}(t)\right)\right], E\left[\widetilde{\mathbf{q}}^*\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}, \mathbf{S}(t)\right)\right]\right) \\ &< d\left(\widetilde{\widehat{\mathbf{q}}^{\mathbf{1}}}, \widetilde{\widehat{\mathbf{q}}^{\mathbf{2}}}\right). \end{aligned}$$

Thus, we have a contradiction. Hence we can have at most one solution. In light of the first intermediate result, we conclude that (20) has a unique solution which in turn implies the main result. ■

Appendix C
Proof of Lemma 4

Proof: We prove the result by separately considering the cases where S1 and S2 of Assumption 2 hold.

Suppose S1 holds. We consider the update equation (13) as a stochastic approximation update equation, and use Theorem 2.1 of Chapter 5 from [14] to prove the result under S1.

In the following, we show that all the assumptions required to use the theorem are satisfied. The variables corresponding to $\boldsymbol{\theta}_t$, $\mathbf{Y}_t$ and ϵ_t associated with Theorem 2.1 are listed next: $\boldsymbol{\theta}_t = \widehat{\mathbf{q}}^*(t)$, $(Y_t)_i = \lambda_i\left(q_i^*\left(\widehat{\mathbf{q}}^*(t), \mathbf{S}(t+1)\right) - \widehat{q}_i^*(t)\right)\ \forall\ i \in \mathcal{N}$, and $\epsilon_t = \frac{1}{t+\lambda_{\max}}$, for each t.

(A.2.1) is satisfied since $\widehat{\mathbf{q}}^*(t)$ is updated using (13), and $\mathbf{q}^*(\widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t)) \in \mathcal{Q}^N$ for any t. (A.2.2) holds by choosing

$$\left(\bar{\mathbf{g}}(\widehat{\mathbf{q}}^*(t))\right)_i = \lambda_i\left(E\left[q_i^*\left(\widehat{\mathbf{q}}^*(t), \mathbf{S}(t+1)\right)\right] - \widehat{q}_i^*(t)\right),\ \beta_t = 0\,\forall\, t.$$

The measurability of $\bar{\mathbf{g}}$ essentially follows from its continuity discussed next. (A.2.3) holds since $\bar{\mathbf{g}}(\widehat{\mathbf{q}})$ is a continuous function of $\widehat{\mathbf{q}}$ (from Lemma 2 (c)). (A.2.4) and (A.2.5) hold since $\epsilon_t = 1/\left(t+\lambda_{\max}\right)\ \forall\ n$ and $\beta_t = 0\ \forall\ t$.

From Lemma 2 (d), $\forall\ i \in \mathcal{N}$

$$\left(\bar{\mathbf{g}}(\widehat{\mathbf{q}})\right)_i = -\left(\nabla E\left[h\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right]\right)_i.$$

From Lemma 2 (c) and (d), we have that $E\left[h\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right]$ is a continuously differentiable function. Hence (A.2.6) holds.

Thus conditions (A.2.1)-(A.2.6) are satisfied. From Theorem 2.1 and Lemma 3 (iii), we can conclude that on almost all sample paths, $(\widehat{\mathbf{q}}^*(t))_t$ converges to the unique solution of

$$E\left[\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right] = \widehat{\mathbf{q}}.$$

Now, suppose that S2 holds. We again consider (13) as a stochastic approximation update equation, and use Theorem 1.1 of Chapter 6 from [14] to prove the result under S2.

In the following, we show that all the assumptions required to use the theorem are satisfied. The following variables and functions $\boldsymbol{\theta}_t$, $\boldsymbol{\xi}_t$, $\mathbf{Y}_t$, ϵ_t, sigma algebras $\mathcal{F}_t$, the function $\mathbf{g}$, $\boldsymbol{\beta}_t$, $\delta\mathbf{M}_t$ and $\boldsymbol{Z}_t$, appearing in the exposition of Theorem 2.1, correspond to the following variables and functions in our problem setting: $\boldsymbol{\theta}_t = \widehat{\mathbf{q}}^*(t)$, $\boldsymbol{\xi}_t = \mathbf{S}(t+1)$, $(Y_t)_i = \lambda_i\left(q_i^*\left(\widehat{\mathbf{q}}^*(t), \mathbf{S}(t+1)\right) - \widehat{q}_i^*(t)\right)\ \forall\ i \in \mathcal{N}$, $\epsilon_t = \frac{1}{t+\lambda_{\max}}$ for each t, $\mathcal{F}_t$ is such that $(\boldsymbol{\theta}_0, \mathbf{Y}_{i-1}, \boldsymbol{\xi}_i, i \leq t)$ is $\mathcal{F}_t$-measurable, $\left(g\left(\widehat{\mathbf{q}}, \mathbf{s}\right)\right)_i = \lambda_i\left(q_i^*\left(\widehat{\mathbf{q}}, \mathbf{s}\right) - \widehat{q}_i^*\right)\ \forall\ i \in \mathcal{N}$, $\boldsymbol{\beta}_t = \mathbf{0}$ for each t, $\delta\mathbf{M}_t = \mathbf{0}$ for each t, and $\boldsymbol{Z}_t = \mathbf{0}$ for each t.

The Equation 5.1.1 in [14] is satisfied due to our choice of ϵ_t. (A.1.1) is satisfied since $\widehat{\mathbf{q}}^*(t)$ is updated using (13), and $\mathbf{q}^*(\widehat{\mathbf{q}}^*(t-1), \mathbf{s}(t)) \in \mathcal{Q}^N$ for any t. (A.1.2) holds since $g\left(\widehat{\mathbf{q}}, \mathbf{s}\right)$ is a continuous function of $\widehat{\mathbf{q}}$ for any value of $\mathbf{s}$ (from Lemma 2 (c)). (A.1.3) holds since we can choose the function $\bar{\mathbf{g}}$ as follows

$$\left(\bar{\mathbf{g}}\left(\widehat{\mathbf{q}}^*(t)\right)\right)_i = \lambda_i\left(E\left[q_i^*\left(\widehat{\mathbf{q}}^*(t), \mathbf{S}(t+1)\right)\right] - \widehat{q}_i^*(t)\right).$$

From Section 6.2 of [14], if ϵ_t does not go to zero faster than the order of $\frac{1}{\sqrt{t}}$, for (A.1.3) to hold we only need to show that the strong law of large numbers holds for $\left(g\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right) - \bar{\mathbf{g}}\left(\widehat{\mathbf{q}}\right)\right)_t$ for any $\widehat{\mathbf{q}}$. Strong law of large numbers holds since $(\mathbf{S}(t))_t$ is a stationary ergodic random process. (A.1.4) and (A.1.5) hold since $\boldsymbol{\beta}_t = \mathbf{0}$ and $\delta \mathbf{M}_t = \mathbf{0}$ for each t. For checking (A.1.6) and (A.1.7), we use some sufficient conditions discussed in [14] following the theorem. (A.1.6) holds since $g(\widehat{\mathbf{q}}, \mathbf{s})$ is bounded. (A.1.7) holds since $g(\widehat{\mathbf{q}}, \mathbf{s})$ is continuous for any $\mathbf{s}$, and since $|\mathcal{S}|$ is finite. Also, (A5.2.6) holds which can be shown in a manner similar to that done under assumption S1. Thus, from Theorem 1.1 and Lemma 3, we can conclude that on almost all sample paths, $(\widehat{\mathbf{q}}^*(t))_t$ converges to the unique solution of

$$E\left[\mathbf{q}^*\left(\widehat{\mathbf{q}}, \mathbf{S}(t)\right)\right] = \widehat{\mathbf{q}}.$$

■