



**HAL**  
open science

## Control of Multiple Remote Servers for Quality-Fair Delivery of Multimedia Contents

Nesrine Changuel, Bessem Sayadi, Michel Kieffer

► **To cite this version:**

Nesrine Changuel, Bessem Sayadi, Michel Kieffer. Control of Multiple Remote Servers for Quality-Fair Delivery of Multimedia Contents. *IEEE Journal on Selected Areas in Communications*, 2014, 32 (4), pp.746-759. 10.1109/JSAC.2014.140407 . hal-00935867

**HAL Id: hal-00935867**

**<https://centralesupelec.hal.science/hal-00935867>**

Submitted on 24 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

# Control of Multiple Remote Servers for Quality-Fair Delivery of Multimedia Contents

Nesrine Changuel\*, Bessem Sayadi\*, Michel Kieffer<sup>o,+,\dagger</sup>

\* Alcatel-Lucent - Bell-Labs, 91620 Nozay

Email: {nesrine.changuel, bessem.sayadi}@alcatel-lucent.com

<sup>o</sup> L2S, CNRS - SUPELEC - Univ Paris-Sud, 91192 Gif-sur-Yvette

<sup>+</sup> partly on leave at LTCI - CNRS - Telecom Paris-Tech, 75014 Paris

<sup>\dagger</sup> Institut Universitaire de France, 75005 Paris

Email: kieffer@lss.supelec.fr

**Abstract**—This paper proposes a control scheme for the quality-fair delivery of several encoded video streams to mobile users sharing a common wireless resource. Video quality fairness, as well as similar delivery delays are targeted among streams. The proposed controller is implemented within some aggregator located near the bottleneck of the network. The transmission rate among streams is adapted based on the quality of the already encoded and buffered packets in the aggregator. Encoding rate targets are evaluated by the aggregator and fed back to each remote video server (fully centralized solution), or directly evaluated by each server in a distributed way (partially distributed solution). Each encoding rate target is adjusted for each stream independently based on the corresponding buffer level or buffering delay in the aggregator. Communication delays between the servers and the aggregator are taken into account.

The transmission and encoding rate control problems are studied with a control-theoretic perspective. The system is described with a multi-input multi-output model. Proportional Integral (PI) controllers are used to adjust the video quality and control the aggregator buffer levels. The system equilibrium and stability properties are studied. This provides guidelines for choosing the parameters of the PI controllers.

Experimental results show the convergence of the proposed control system and demonstrate the improvement in video quality fairness compared to a classical transmission rate fair streaming solution and to a utility max-min fair approach<sup>1</sup>.

**Index Terms**—Command and control systems; Decentralized control; Multimedia communication; Quality of service; Stability analysis.

## I. INTRODUCTION

With the development of wireless networks and widespread of smartphones, delivery of compressed videos (video-on-demand or mobile TV broadcast services) to mobile users is increasing rapidly. This trend is likely to continue in the coming years [1]. To satisfy the related increasing demand for resources, operators have to expand their network capacity with as limited as possible infrastructure investments. In parallel, they have to optimize the way multimedia contents are

delivered to users while satisfying *application-layer* quality-of-service (QoS) constraints, which are more challenging to address than traditional *network-layer* QoS constraints.

Delivered videos have a large variety of quality-rate characteristics, whatever the considered quality metric, *e.g.*, Peak Signal-to-Noise Ratio (PSNR), Structural SIMilarity (SSIM) [2], *etc.* These characteristics are time-varying and depend on the content of the videos. Provisioning some constant transmission rate to mobile users for video delivery is in general inappropriate. If videos are encoded at a constant bitrate, the quality may fluctuate with the variations of the characteristics of the content. If they are encoded at a variable bitrate, targeting a constant quality, buffering delays may fluctuate significantly.

This paper proposes a controller for the quality-fair delivery of several encoded video streams to mobile users sharing a common wireless resource. Video-on-demand or multi-cast/broadcast transmission are typical applications for this scenario. Video encoding rate adaptation and wireless resource allocation are performed jointly within some Media Aware Network Element (MANE) using feedback control loops. The aim is to provide users with encoded videos of similar quality and with controlled delivery delay, without exchanging information between remotely located video servers.

### A. Related work

When controlling the parallel delivery of several video streams, their rate-distortion (R-D) trade-off may be adjusted by selectively discarding frames as in [3], [4] or via an adaptation of their encoding parameters as in [5], [6]. With scalable video encoders, such as H.264/SVC, the R-D trade-off may be adjusted via packet filtering [7], [8]. In this case, the control parameter is the number of transmitted enhancement layers for each frame.

If several video streams are transmitted to different users in a dedicated broadcast channel with limited capacity, a blind source rate allocation could lead to unacceptable quality for high-complexity video contents compared to low-complexity ones. Therefore, providing fairness is an important issue that must be addressed.

<sup>1</sup>Parts of this work have been presented at ACM Multimedia conference, 2012. This work has been partly supported by ANR ARSSO project, contract number ANR-09-VERS-019-02 and by ANR project LimCoS, contract number ANR-12-BS03-005-01.

Video quality fairness among encoded streams may be obtained by sharing quality information, or R-D characteristics via a central controller providing to each server quality or rate targets, as in [9], [10]. This technique enables the encoders to adjust their bit rate or to drop frames or quality layers depending on the complexity of the videos and on the available transmission rate.

Control-theoretic approaches have been considered to address the problem of rate control in the context of video streaming, see, *e.g.*, [11], [12], [13], [14], [15]. In [11], a real-time rate control based on a Proportional Integral Derivative (PID) controller is proposed for a single video stream. The main idea is to determine the encoding rate per frame based on the buffer level to maximize video quality and minimize quality variations over frames. In [12], a flow control mechanism with active queue management and a proportional controller is considered. A flow control is used to reduce the buffer size and avoid buffer overflow and underflow. The flow control mechanism is shown to be stable for small buffer sizes and non-negligible round-trip times. In [13], a rate allocation algorithm, performed at the Group of Picture (GoP) level, is performed at the sender to maximize the visual quality according to the overall loss and the receiver buffer occupancy. This target is achieved using a Proportional Integral (PI) controller in charge of determining the transmission rate to drain the buffers. Later, [16] introduces a rate controller that uses different bit allocation strategies for Intra and Inter frames. A PID controller is adopted to minimize the deviations between the target and the current buffer level. Buffer management is performed at the bit level and delivery delay is not considered. Moreover, [11], [12], [13], [16] address single flow transmission.

For multi-video streaming, in [17], a distributed utility max-min flow control in the presence of round-trip delays is proposed. The distributed link algorithm performs utility max-min bandwidth sharing while controlling the link buffer occupancy around a target value at the cost of link under utilization using a PID controller. Stability analysis in case of a single bottleneck and homogeneous delay is conducted. In [9], a content-aware distortion-fair video delivery scheme is proposed to deliver video streams based on the characteristics of video frames. It provides a max-min distortion fair resource sharing among video streams. The system uses temporal prediction structure of the video sequences with a frame drop strategy based on the frame importance to guide resource allocation. The proposed scheme is for video on-demand services, where the rate and the importance of each frame are assumed calculated in advance. A proportional controller is considered in [14] to stabilize the received video quality as well as the bottleneck link queue for both homogeneous and heterogeneous video contents. A PI controller is considered in [15]. Robustness and stability properties are studied. In [14] and [15], the rate control is performed in a centralized way, exploiting the rate and distortion characteristics of the considered video flows to determine the encoding rate for the next frame of each flow. In [18], a cross-layer optimization framework for scalable video delivery over OFDMA wireless systems is proposed, aiming at maximizing the sum of the

achievable rates while minimizing the distortion difference among multiple videos. The optimization problem is described by a Lagrangian constrained sum-rate maximization to achieve distortion fairness among users. However the communication delay between the control block and the controlled servers is not addressed.

Remotely implemented control laws are also considered in [19] leading to the problem of stabilizing an open-loop system with time-varying delay. The problem of remote stabilization via communication networks is considered with an explicit use of the average network dynamics and an estimation of the average delay in the control law. The control law does not address video transmission issues, so no quality constraint on the transmitted data is considered.

The commercial products described in [20], [21] propose statistical multiplexing systems enabling encoders to adapt their outputs to the available channel rate. Connecting encoders and multiplexers via a switched IP network allows collocated and distributed encoders to be part of the multiplexing system. Nevertheless, in these solutions, quality fairness constraints between programs appear not to be considered among the video quality constraints.

## B. Main contributions

In this paper, we propose a control system to perform jointly (i) encoding rate control of spatially spread video servers without information exchange between them and (ii) transmission rate control of the encoded streams through some bottleneck link. A MANE, located near the bottleneck link derives the average video quality of the data stored in dedicated buffers fed by the remote servers. This average video quality is compared by each individual transmission rate controller to the quality of its video flow to adjust the draining rate of the corresponding buffer (first control input). For that purpose, programs with low quality are drained faster than programs with high quality. Dedicated encoding rate controllers observe the buffer levels to adjust the video encoding rates (second control input). The encoding rate control targets a similar buffer level for all programs. The buffer level in bits or the buffering delay can be adjusted via an adaptation of the video encoding rates, *e.g.*, by scalability layer filtering when a scalable video coder is involved.

In a *fully centralized* version of the controller, the MANE is in charge of sending the encoding rate target to each video server. In a *partly distributed* version, only the individual buffer level discrepancies are transmitted to the servers, which are then in charge of computing their own encoding rate target. Communication delays between the MANE and the servers are considered in both directions. A discrete-time state-space representation of the system is introduced. The buffer level (in bits) or the buffering delay has to be controlled and quality fairness between video streams has to be obtained. For that purpose, feedback loops involving PI controllers are considered. The quality fairness constraint among streams leads to a coupling of the state equations related to the control of the delivery of each stream. The system equilibrium and stability properties are studied. This provides guidelines for

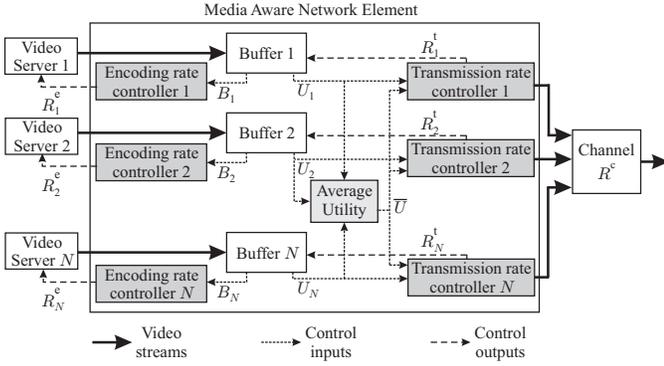


Figure 1. Structure of the proposed quality-fair video delivery system (fully centralized version).

choosing the parameters of the PI controllers. This paper extends preliminary results obtained in [22], where control of the buffer level in bits is not addressed and the communication delay between the MANE and the servers is not explicitly considered.

Section II introduces the considered system and the constraints that have to be satisfied. The proposed solution is described in Section III. The required hypotheses are listed and a discrete-time state-space representation of the system is provided, emphasizing the coupling between equations induced by the fairness constraint among video streams. The equilibrium and stability analyses are performed in Section IV. Finally, a typical application context is described in Section V and experimental results are detailed. Robustness of the proposed control system to variations of the channel rate and of the number of video streams is shown.

## II. SYSTEM DESCRIPTION

Consider a communication system in which  $N$  encoded video streams provided by  $N$  remote servers arrive to some network bottleneck where they have to share a communication channel providing a total transmission rate  $R^c$ , see Figure 1. The servers deliver encoded Video Units (VUs) representing a single picture or a Group of Pictures (GoP). All VUs are assumed to be of the same duration  $T$  and the frame rate  $F$  is assumed constant over time and identical for all streams. Time is slotted and the  $j$ -th time index represents the time interval  $[jT, (j+1)T)$ .

In the proposed system, a Media Aware Network Element (MANE) is located close to the bottleneck of the network, see Figure 1. The MANE aims at providing the receivers with video streams of similar (objective or subjective) quality and with similar delivery delays. For that purpose, two feedback loops are considered to control (i) the encoding rate and (ii) the transmission rate of each video stream, see Figure 2.

Encoded and packetized VUs are temporarily stored in dedicated buffers in the MANE. We assume that a *utility*  $U_i(j)$  measures the quality of each VU  $j$  for each stream  $i$  (in terms of PSNR, SSIM, or any other video quality metric [23]). The MANE has access to  $U_i(j)$ , stored, e.g., in the packet headers. The transmission rate controllers are in charge of choosing the draining rates from each buffer so that all utilities within the  $N$

buffers are as close as possible. The encoding rate controllers are in charge of choosing the video encoding rates so that the buffer levels in the MANE are adjusted around some reference level  $B_0$  in bits or reference delay  $\tau_0$  in seconds. This control is performed at each discrete time index.

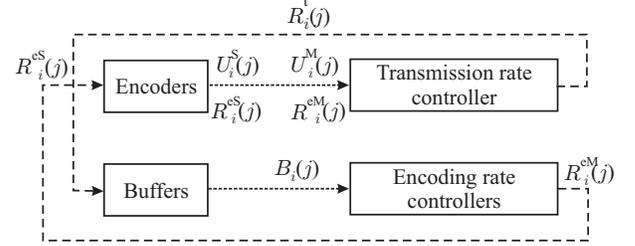


Figure 2. Feedback control loops in the proposed quality-fair video delivery system.

The MANE evaluates the average utility and each transmission rate controller allocates more rate to streams with a utility below average. The buffers of these streams are drained faster than those with utility above average. This control, done within the MANE, is thus *centralized*. The discrepancy of each buffer level in bits (respectively delay in seconds) with respect to the reference level  $B_0$  (respectively delay  $\tau_0$ ) is processed individually by an encoding rate controller. The encoding rate for the next VU of each video stream is then evaluated to regulate the buffer level around the reference level. The encoding rate may be evaluated at the MANE, in which case, the control is fully *centralized* and the video encoders/servers receive only the evaluated target bit rate. Alternatively, the encoding rates may be evaluated in a *decentralized* way at each video encoder/server. For the latter case, the MANE only feeds back to each server the buffer level (respectively delay) discrepancy. This solution requires all video encoders/servers to host individually an encoding rate controller, which is mainly possible for managed video servers. In this paper, the encoding rate target is evaluated within the MANE, but the encoding parameters are evaluated in a *distributed* way by each server [24].

The interaction of both control loops (transmission rate control and encoding rate control) allows getting a quality-fair video delivery. Videos with a quality below average have a buffer in the MANE that is drained faster, and thus is likely to be below  $B_0$  or  $\tau_0$ . The encoding rate of such streams is then increased, to improve their quality.

Feedback delays between the MANE and the video servers are considered. They correspond to the delays introduced when the servers deliver encoded packets to the MANE and when the MANE feeds back signalization to carry encoding rate targets to the servers. To account for the delayed utility available at the MANE and the delayed encoding rate targets sent by the MANE to the servers, two state variables representing the delayed encoding rates and the delayed utilities are introduced

$$R_i^{eS}(j) = R_i^{eM}(j - \delta_1), \quad (1)$$

and

$$U_i^M(j) = U_i^S(j - \delta_2), \quad (2)$$

where  $R_i^M(j)$  is the encoding rate evaluated at the MANE for the  $j$ -th VU of program  $i$ . This encoding rate target reaches the server with some delay  $\delta_1$ , where it is denoted by  $R_i^S(j)$ . On the other hand, the server sends the utility of the  $j$ -th VU of program  $i$ , denoted by  $U_i^S(j)$ . It arrives at the MANE with some delay  $\delta_2$  and is denoted  $U_i^M(j)$ , see Figure 2. The M superscript refers to the information available at the MANE and the S superscript refers to the information available at the server for both rate and utility. In what follows, the stability of both control loops is studied.

### III. STATE-SPACE REPRESENTATION

A state-space representation of the system presented in Section II is introduced to study its stability. Several additional assumptions are needed to get a tractable representation.

#### A. Assumptions

1) *Feedback delay*: In what follows, a VU represents a GoP. All encoded VUs processed by the video servers during the  $(j-1)$ -th time slot  $[(j-1)T, jT)$  are assumed to have reached the MANE during the  $j$ -th time slot  $[jT, (j+1)T)$ . The encoding rate targets sent by the MANE to the servers  $R_i^e(j)$  at the beginning of the  $j$ -th time slot are assumed to have reached all video servers at the beginning of the  $(j+1)$ -th time slot. This rate is used to encode the  $(j+1)$ -th VU which will be placed in the buffer in the MANE during the  $(j+2)$ -th time slot, *etc.*, see Figure 3 when  $\delta_1 = \delta_2 = 1$ .

The transmission delays between the MANE and the servers may vary. The previous assumptions allow to cope with forward and backward delays upper bounded by  $T$ . This is reasonable, since the network transmission and buffering delays are of the order of tens of milliseconds, provided that it is not too congested. This is less than the duration of VUs when they represent GoPs (typically 0.25 s to 1 s). Following the bounded-delay assumption, during the  $j$ -th time slot, the MANE has only access to the utilities  $U_i^S(j-2)$ ,  $i = 1, \dots, N$  of the  $(j-2)$ -th encoded VUs.

In the rest of the paper the superscripts S and M are omitted. Then  $U_i(j)$  is the utility of the  $j$ -th VU of the  $i$ -th stream encoded during time slot  $j$ , transmitted to the MANE during time slot  $j+1$ , and fully buffered in the MANE at the beginning of time slot  $j+2$ .  $R_i^e(j)$  is the encoding rate target evaluated by the MANE during the  $j$ -th time slot.  $R_i^e(j)$  reaches the server at the beginning of time slot  $j+1$ , see Figure 3.

2) *Source model*: The following parametric rate-utility model is used to describe the evolution of the utility  $U_i(j)$  as a function of the rate  $R_i^e(j)$  used to encode the  $j$ -th VU of the  $i$ -th stream

$$U_i(j) = f(\mathbf{a}_i(j), R_i^e(j)), \quad (3)$$

where  $\mathbf{a}_i(j) \in \mathcal{A} \subset \mathbb{R}^{N_a}$  is a time-varying and program-dependent parameter vector. Note that this model is only used to define the controller parameters so that the system is stable. Once these parameters are set, the rate-utility model is no more needed. For all values of  $\mathbf{a}$  belonging to the set of admissible parameter values  $\mathcal{A}$ ,  $f(\mathbf{a}, R)$  is assumed to be a continuous

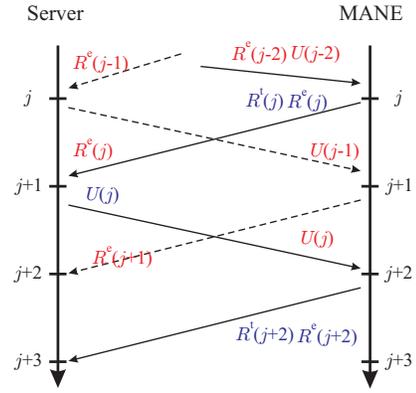


Figure 3. Communication delays between a video server and the MANE.

and strictly increasing function of  $R$ , with  $f(\mathbf{a}, 0) = 0$ . The variation with time of  $\mathbf{a}_i(j)$  is described by

$$\mathbf{a}_i(j+1) = \mathbf{a}_i(j) + \delta \mathbf{a}_i(j), \quad (4)$$

where  $\delta \mathbf{a}_i(j)$  represents the uncontrolled variations of the source characteristics.

The model (3) may represent the variation with the encoding rate of the SNR, the PSNR, the SSIM, or any other strictly increasing quality metric.

3) *Buffer model*: As introduced in Section II, the MANE contains dedicated buffers for each of the  $N$  encoded video streams. The evolution of the level in bits of the  $i$ -th buffer between time slot  $j$  and  $j+1$  is

$$B_i(j+1) = B_i(j) + (R_i^e(j-2) - R_i^l(j))T, \quad (5)$$

where  $R_i^l(j)$  is the transmission rate for the  $i$ -th stream and  $R_i^e(j-2)$  is the encoding rate of the  $(j-2)$ -th VU evaluated at the MANE at time slot  $j$ . In (5),  $R_i^e(j-2)$  accounts for the communication delay between the server and the MANE, see Figure 3.

Buffers are controlled in two different ways. A control of the buffer level in bits  $B_i(j)$  maintains an averaged level in bits to prevent buffer overflow and underflow. This is appropriate for applications with buffers of limited size. A control of the buffering delay helps adjusting the end-to-end delivery delay. This type of control is better suited for delay-sensitive applications. Buffering delay control within the MANE also allows implicitly controlling the buffering delay at the client for live and broadcast applications. This is due to the fact that the end-to-end delay between the transmission of a VU by the server and its playback by the client is constant over time for live video<sup>2</sup>.

Let  $h_i(j)$  be the number of VUs in the  $i$ -th MANE buffer at time  $j$ , the corresponding buffering delay is

$$\tau_i(j) = h_i(j)T. \quad (6)$$

Assume that packets containing encoded VUs are segmented to allow a fine granularity of the draining rate. Then  $h_i(j)$  becomes quite difficult to evaluate accurately and directly within the MANE using only information stored in packet headers.

<sup>2</sup>Some periodic feedback may nevertheless be useful to verify that the system actually behaves nominally.

The corresponding buffering delay is then approximatively evaluated as

$$\tau_i(j) = \frac{B_i(j)}{\bar{R}_i^e(j)}, \quad (7)$$

where

$$\begin{aligned} \bar{R}_i^e(j) &= \frac{1}{h_i(j)} \sum_{\ell=2}^{\lfloor h_i(j) \rfloor} R_i^e(j - \ell) \\ &+ \frac{1}{h_i(j)} R_i^e(j - \lceil h_i(j) \rceil) (h_i(j) - \lfloor h_i(j) \rfloor) \end{aligned} \quad (8)$$

is the average encoding rate of the VUs stored in the  $i$ -th buffer at time  $j$  and  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  denote rounding towards  $-\infty$  and  $+\infty$ . Since (8) still requires the availability of  $h_i(j)$ , we propose to estimate it as follows

$$\begin{aligned} \tilde{R}_i^e(j) &= R_i^e(j), \text{ if } j \leq 2, \\ \tilde{R}_i^e(j) &= \alpha R_i^e(j - 2) + (1 - \alpha) \tilde{R}_i^e(j - 1), \text{ if } j > 2, \end{aligned} \quad (9)$$

where  $0 < \alpha < 1$  is some tuning parameter. Then, one gets an estimate of the buffering delay (6) using (9)

$$\tilde{\tau}_i(j) = \frac{B_i(j)}{\tilde{R}_i^e(j)}. \quad (10)$$

## B. Rate controllers

$N$  coordinated transmission rate controllers and  $N$  (possibly distributed) encoding rate controllers are considered in the video delivery system described in Section II. A PI control of the transmission rate is performed. At time  $j$ , each PI controller takes as input the average utility evaluated by the MANE and the utility  $U_i(j - 2)$  of the VU  $j - 2$  of the controlled stream. PI controllers are also used to evaluate the target encoding rate of each program in order to regulate the buffer level around  $B_0$  or the buffering delay around  $\tau_0$ .

1) *Transmission rate control*: At time  $j$ , the available channel rate  $R^c$  is shared between the  $N$  video streams. The delayed utility of the VU available at the MANE at time  $j$  for the  $i$ -th stream is

$$U_i^{dd}(j) = U_i^d(j - 1) = U_i(j - 2) \quad (11)$$

and the utility discrepancy  $\delta U_i^{dd}(j)$  with the *average* utility  $\bar{U}(j)$  over the  $N$  programs at time  $j$  is

$$\delta U_i^{dd}(j) = \frac{1}{N} \sum_{\ell=1}^N (U_\ell^{dd}(j) - U_i^{dd}(j)) = \bar{U}(j) - U_i^{dd}(j). \quad (12)$$

The PI transmission rate controller for the  $i$ -th program uses  $\delta U_i^{dd}(j)$  to evaluate the transmission rate allocated to each video stream

$$R_i^t(j) = R_0 + (K_P^t + K_I^t) \delta U_i^{dd}(j) + K_I^t \phi_i(j), \quad (13)$$

where  $K_P^t$  and  $K_I^t$  are the proportional and integral correction gains. All rates are evaluated with respect to  $R_0 = R^c/N$ , the average encoding rate per stream that would be used when the  $N$  streams represent the same encoded video. In (13),  $\phi_i(j)$  is the cumulated utility discrepancy (used for the integral term) evaluated as

$$\begin{aligned} \phi_i(j) &= 0, \text{ if } j \leq 2, \\ \phi_i(j + 1) &= \phi_i(j) + \delta U_i^{dd}(j), \text{ if } j > 2. \end{aligned} \quad (14)$$

$R_i^t(j)$  represents the draining rate of the  $i$ -th MANE buffer at time  $j$ . One may easily verify that

$$\sum_{i=1}^N R_i^t(j) = R^c. \quad (15)$$

According to (13), in a first approximation, more (resp. less) transmission rate is allocated to programs with a utility below (resp. above) average.

2) *Encoding rate control*: The encoding rate control is performed independently for each video stream. This allows a distributed implementation of this part of the global controller.

For the control of the buffer level or of the buffering delay, the buffer level (5) is needed. Let  $R_i^{edd}(j)$  be the delayed encoding rate of the VU reaching the MANE at time  $j$  for the  $i$ -th program

$$R_i^{edd}(j) = R_i^{ed}(j - 1) = R_i^e(j - 2). \quad (16)$$

Using (5) and (16), one gets

$$B_i(j + 1) = B_i(j) + (R_i^{edd}(j) - R_i^t(j)) T. \quad (17)$$

In case of buffer level control, the encoding rate for the  $j$ -th VU of each video program is controlled to limit the deviations of  $B_i(j)$  from the reference level  $B_0$ . At time  $j$ , the discrepancy  $\delta B_i(j)$  between  $B_i(j)$  and  $B_0$  is

$$\delta B_i(j) = B_i(j) - B_0. \quad (18)$$

In case of buffering delay control, the encoding rate for the  $j$ -th VU of each video program is controlled to limit the deviations of  $\tilde{\tau}_i(j)$  from the reference level  $\tau_0$ . At time  $j$ , the discrepancy  $\delta \tau_i(j)$  between  $\tilde{\tau}_i(j)$  and  $\tau_0$  is

$$\delta \tau_i(j) = \tilde{\tau}_i(j) - \tau_0 = \left( \frac{B_i(j)}{\tilde{R}_i^{edd}(j)} - \tau_0 \right). \quad (19)$$

Buffers with positive  $\delta B_i(j)$  or  $\delta \tau_i(j)$  contain more than  $B_0$  bits or  $\tau_0$  seconds of encoded videos. The encoding rate  $R_i^e(j)$  has thus to be decreased. This part of the control process is very similar to back-pressure algorithms [25].

$R_i^e(j)$  is evaluated as the output of a PI controller

$$R_i^e(j) = R_0 - \frac{K_P^e + K_I^e}{T} \delta \tau_i(j) - \frac{K_I^e}{T} \Pi_i(j). \quad (20)$$

where  $K_P^e$  and  $K_I^e$  are the proportional and integral gains and  $\Pi_i(j)$  is the cumulated buffer discrepancy in seconds evaluated as

$$\begin{aligned} \Pi_i(j) &= 0, \text{ if } j \leq 3, \\ \Pi_i(j + 1) &= \Pi_i(j) + \delta \tau_i(j), \text{ if } j > 3. \end{aligned} \quad (21)$$

For a control of the buffer level,  $\delta \tau_i(j)$  is replaced by  $\delta B_i(j)$  in (20) and (21).

Taking into account the communication delay between the MANE and the server, the encoding rate target  $R_i^e(j)$  evaluated at the MANE at time  $j$  reaches the video server at time  $j + 1$ . Thus,  $R_i^e(j)$  represents the encoding rate for the  $j + 1$ -th VU. The encoding rate increases (resp. decreases) when the buffer is below (resp. above) its reference level. The sum of the encoding rates is not necessarily equal to  $R^c$ . This allows to compensate for the variations of the video characteristics.

Considering simultaneously (5), (13), and (20), one sees that buffers corresponding to programs producing video with lower utility than average are drained faster. As a consequence, the encoding rate allowed to encode the next VU of such programs is increased, potentially increasing the utility.

### C. State-space representation

The state-space representation facilitates the study of the system equilibrium and stability properties. Two representations are considered, depending on whether the buffer level or the buffering delay is controlled.

In the case of a control of the buffering delay, combining (3), (4), (11), (13), (14), (17), (20), and (21) leads to the following discrete-time nonlinear state-space representation for the  $i$ -th video stream,  $i = 1, \dots, N$

$$\mathbf{a}_i(j+1) = \mathbf{a}_i(j) + \delta \mathbf{a}_i(j) \quad (22a)$$

$$\mathbf{a}_i^d(j+1) = \mathbf{a}_i(j) \quad (22b)$$

$$\phi_i(j+1) = \phi_i(j) + \frac{1}{N} \sum_{k=1}^N U_k^{\text{dd}}(j) - U_i^{\text{dd}}(j) \quad (22c)$$

$$\Pi_i^\tau(j+1) = \Pi_i^\tau(j) + \left( \frac{B_i(j)}{\tilde{R}_i^e(j)} - \tau_0 \right) \quad (22d)$$

$$\tilde{R}_i^e(j+1) = \alpha R_i^{\text{edd}}(j) + (1 - \alpha) \tilde{R}_i^e(j) \quad (22e)$$

$$R_i^{\text{ed}}(j+1) = R_0 - \frac{K_P^{\text{er}} + K_I^{\text{er}}}{T} \left( \frac{B_i(j)}{\tilde{R}_i^e(j)} - \tau_0 \right) - \frac{K_I^{\text{er}} \tau}{T} \Pi_i(j) \quad (22f)$$

$$R_i^{\text{edd}}(j+1) = R_i^{\text{ed}}(j) \quad (22g)$$

$$U_i^{\text{dd}}(j+1) = f(\mathbf{a}_i^d(j), R_i^{\text{ed}}(j)) \quad (22h)$$

$$B_i(j+1) = B_i(j) + R_i^{\text{edd}}(j)T - R_0T - \left( (K_P^{\text{t}} + K_I^{\text{t}}) \left( \frac{1}{N} \sum_{k=1}^N U_k^{\text{dd}}(j) - U^{\text{dd}}(j) \right) + K_I^{\text{t}} \phi_i(j) \right) T \quad (22i)$$

where  $\mathbf{a}_i^d(j)$  is the delayed video characteristic vector of the  $(j-1)$ -th VU. The utilities  $U_k^{\text{dd}}(j)$  of all video streams appear in (22), leading to a coupling of the state-space representations related to the control of the individual video streams.

When buffer level control is addressed,  $B_i(j)/\tilde{R}_i^e(j)$  is replaced by  $B_i(j)$  in (22d) and (22f), and the state (22e) does not appear anymore.

In the remainder of the paper, the subscript  $b$  is for buffer level control and the subscript  $\tau$  is for buffering delay control.

## IV. EQUILIBRIUM AND STABILITY

The steady-state behavior and the stability of the video delivery system described by (22) for buffering delay control as well as the simpler system for buffer level control are studied. Due to the coupling between controllers induced by the constraint that the discrepancy between the average utility and the utility of each program has to be as small as possible, both characterizations have to be done on the whole system. In the rest of this section, we derive the equilibrium and perform the stability analysis for the control system where buffering

delays are controlled. The technique is similar when for the control of the buffer level.

### A. Equilibrium analysis

The system reaches an equilibrium when all terms on the left of the state-space representation (22) do not change with time. This leads to a system of  $(N_a + 6) \times N$  equations with  $(N_a + 6) \times N$  unknowns.

$$\begin{cases} \delta \mathbf{a}_i^{\text{eq}} &= 0 \\ U_i^{\text{eq}} &= \frac{1}{N} \sum_{k=1}^N U_k^{\text{eq}} \\ B_i^{\text{eq}} &= \tau_0 \tilde{R}_i^{\text{e,eq}} \\ \tilde{R}_i^{\text{e,eq}} &= R_i^{\text{e,eq}} \\ R_i^{\text{e,eq}} &= R_0 - \frac{K_I^{\text{er}} \tau}{T} \Pi_i^{\tau \text{eq}} \\ U_i^{\text{eq}} &= f(\mathbf{a}_i^{\text{eq}}, R_i^{\text{e,eq}}) \\ R_i^{\text{e,eq}} &= K_I^{\text{t}} \phi_i^{\text{eq}} \end{cases}, \quad (23)$$

The second equation in (23) imposes that  $U_1^{\text{eq}} = \dots = U_N^{\text{eq}} = U^{\text{eq}}$ ; all programs have thus the same utility at equilibrium. Moreover, one has  $\tilde{R}_i^{\text{e,eq}} = R_i^{\text{e,eq}}$  and  $B_i^{\text{eq}} = \tau_0 R_i^{\text{e,eq}}$ , which leads to  $B_i^{\text{eq}}/R_i^{\text{e,eq}} = \tau_0$ , for  $i = 1, \dots, N$  meaning that at equilibrium, the buffering delay is equal to  $\tau_0$  for all streams.

The target encoding rates at equilibrium  $R_i^{\text{e,eq}}$  and the utility  $U^{\text{eq}}$  are obtained as the solution of a system of  $N+1$  equations

$$\begin{cases} f(\mathbf{a}_1^{\text{eq}}, R_1^{\text{e,eq}}) = \dots = f(\mathbf{a}_N^{\text{eq}}, R_N^{\text{e,eq}}) = U^{\text{eq}} \\ \sum_{i=1}^N R_i^{\text{e,eq}} = R^c \end{cases}, \quad (24)$$

depending of the values  $\mathbf{a}_i^{\text{eq}}, i = 1, \dots, N$  of the parameter vector of the rate-utility model. At equilibrium, they are assumed constant in time and well-estimated, see Section V. Since  $f$  is strictly increasing with  $R$ , the rate at equilibrium as a function of  $U^{\text{eq}}$  is

$$R_i^{\text{e,eq}} = f_R^{-1}(\mathbf{a}_i^{\text{eq}}, U^{\text{eq}}), \quad i = 1, \dots, N, \quad (25)$$

with  $f_R^{-1}$  is the inverse of  $f$  seen as a function of  $R$  only. The value of  $U^{\text{eq}}$  is determined from the channel rate constraint

$$\sum_{i=1}^N R_i^{\text{e,eq}} = \sum_{i=1}^N f_R^{-1}(\mathbf{a}_i^{\text{eq}}, U^{\text{eq}}) = R^c. \quad (26)$$

Since  $f(\mathbf{a}, R)$  is a continuous and strictly increasing function of  $R$ ,  $f_R^{-1}(\mathbf{a}, U)$  and  $\sum_{i=1}^N f_R^{-1}(\mathbf{a}_i, U)$  are also continuous and strictly increasing functions of  $U$ , with  $\sum_{i=1}^N f_R^{-1}(\mathbf{a}_i, 0) = 0$ . Provided that

$$\lim_{U \rightarrow \infty} \sum_{i=1}^N f_R^{-1}(\mathbf{a}_i, U) > R^c, \quad (27)$$

(26) admits a unique solution.  $\Pi_i^{\text{eq}}$  and  $\phi_i^{\text{eq}}, i = 1, \dots, N$ , are deduced from (23) and (25), provided that  $K_I^{\text{t}} \neq 0$  and  $K_I^{\text{t}} \neq 0$ .

The equilibrium is thus unique and satisfies the control targets considered in Section II. Similar conclusions can be obtained when the buffer level is controlled.

## B. Linearized model

We study the local stability of the system around an equilibrium point evaluated in Section IV-A. Linearizing (22) around the equilibrium characterized in (23) one gets for  $i = 1, \dots, N$

$$\begin{cases} \Delta \mathbf{a}_i(j+1) = \Delta \mathbf{a}(j) + \delta \mathbf{a}_i(j) \\ \Delta \mathbf{a}_i^d(j+1) = \Delta \mathbf{a}_i(j) \\ \Delta \phi_i(j+1) = \Delta \phi_i(j) + \frac{1}{N} \sum_{k=1}^N \Delta U_k^{\text{dd}}(j) - \Delta U_i^{\text{dd}}(j) \\ \Delta \Pi_i^\tau(j+1) = \Delta \Pi_i^\tau(j) - \frac{1}{R_i^{\text{eq}}} (\tau_0 \Delta \tilde{R}_i^e(j) - \Delta B_i(j)) \\ \Delta \tilde{R}_i^e(j+1) = (1 - \alpha) \Delta \tilde{R}_i^e(j) + \alpha \Delta R_i^{\text{edd}}(j) \\ \Delta R_i^{\text{ed}}(j+1) = \frac{K_p^e + K_I^e}{T} \frac{1}{R_i^{\text{eq}}} (\tau_0 \Delta \tilde{R}_i^e(j) - \Delta B_i(j)) - \frac{K_I^e}{T} \Delta \Pi_i^\tau(j) \\ \Delta R_i^{\text{edd}}(j+1) = \Delta R_i^{\text{ed}}(j) \\ \Delta U_i^{\text{dd}}(j+1) = \frac{\partial f}{\partial \mathbf{a}}(\mathbf{a}_i^{\text{d,eq}}, R_i^{\text{ed,eq}}) \Delta \mathbf{a}_i^d(j) + \frac{\partial f}{\partial R}(\mathbf{a}_i^{\text{d,eq}}, R_i^{\text{ed,eq}}) \Delta R_i^{\text{ed}}(j) \\ \Delta B_i(j+1) = \Delta B_i(j) + \Delta R_i^{\text{edd}}(j) T \\ - \left( (K_p^l + K_I^l) \left( \frac{1}{N} \sum_{k=1}^N \Delta U_k^{\text{dd}}(j) - \Delta U_i^{\text{dd}}(j) \right) + K_I^l \Delta \phi_i(j) \right) T. \end{cases} \quad (28)$$

Consider the  $N \times (N \times N^a)$  block diagonal matrix

$$\Xi = \text{diag} \left( \frac{\partial f}{\partial \mathbf{a}^T}(\mathbf{a}_1^{\text{d,eq}}, R_1^{\text{e,eq}}), \dots, \frac{\partial f}{\partial \mathbf{a}^T}(\mathbf{a}_N^{\text{d,eq}}, R_N^{\text{e,eq}}) \right), \quad (29)$$

gathering the sensitivities with respect to  $\mathbf{a}$  of the rate-utility characteristics of each stream and the  $N \times N$  diagonal matrix

$$\Gamma = \text{diag} \left( \frac{\partial f}{\partial R}(\mathbf{a}_1^{\text{d,eq}}, R_1^{\text{e,eq}}), \dots, \frac{\partial f}{\partial R}(\mathbf{a}_N^{\text{d,eq}}, R_N^{\text{e,eq}}) \right) \quad (30)$$

gathering the sensitivity to  $R$  of the rate-utility characteristics of each stream. Putting all coupled linearized state-space representations (28) together, one gets a linear discrete-time state-space representation

$$\mathbf{x}^\tau(j+1) = \mathbf{A}^\tau \mathbf{x}^\tau(j) + \mathbf{w}(j) \quad (31)$$

with state vector

$$\mathbf{x}^\tau = \left( \Delta \mathbf{a}, \Delta \mathbf{a}^d, \Delta \phi, \Delta \Pi^\tau, \Delta \tilde{R}^e, \Delta R^{\text{ed}}, \Delta R^{\text{edd}}, \Delta U^{\text{dd}}, \Delta B \right)^T \quad (32)$$

and noise input

$$\mathbf{w} = \left( \delta \mathbf{a} \quad \mathbf{0} \quad \dots \quad \mathbf{0} \right)^T, \quad (33)$$

representing the fluctuations of the value of the parameter vector for the rate-utility model. In (32) and (33), boldface letters represent vectors and time indexes have been omitted. For example  $\Delta \mathbf{a}(j)$  is a vector  $N \times N^a$  components and  $\Delta \mathbf{B}(j) = (\Delta B_1(j), \dots, \Delta B_N(j))^T$  is a vector of  $N$  components. From (28) and (31), one deduces

$$\mathbf{A}^\tau = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & -\tau_0 \mathbf{V} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{V} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & (1 - \alpha) \mathbf{I} & \mathbf{0} & \alpha \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\frac{K_p^e}{T} \mathbf{I} & \frac{K_p^e}{T} \tau_0 \mathbf{V} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\frac{K_I^e}{T} \mathbf{V} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Xi & \mathbf{0} & \mathbf{0} & \mathbf{0} & \Gamma & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -K_I^l T \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & T \mathbf{I} & K_I^l T \mathbf{L} & \mathbf{I} \end{pmatrix} \quad (34)$$

with  $\mathbf{V} = \text{diag}(1/R_1^{\text{e,eq}}, \dots, 1/R_N^{\text{e,eq}})$  a diagonal matrix containing the inverse of the encoding rates at equilibrium and  $K^{\tau e} = K_p^{\tau e} + K_I^{\tau e}$ ,  $K^l = K_p^l + K_I^l$ .  $\mathbf{I}$  and  $\mathbf{0}$  are identity and null matrices of appropriate size.

When studying the roots of  $\det(z\mathbf{I} - \mathbf{A}) = 0$ ,  $N_a \times N$  roots at  $z = 1$  are obtained. They correspond to the variations of the

rate-utility parameter vector (4). The matrix  $\Xi$ , representing the sensitivity with respect to  $\mathbf{a}$  of the rate-utility characteristics  $f$ , does not appear in the expressions of  $\mathbf{A}^\tau$ . Only the sensitivity of  $f$  with respect to  $R$ , represented by  $\Gamma$ , impacts the stability around equilibrium. The system stability is also influenced by the encoding rates at equilibrium via  $\mathbf{V}$  and determined by the PI controller gains  $K_p^l$ ,  $K_I^l$ ,  $K_p^e$ , and  $K_I^e$ .

In Section V, values of  $K_p^l$ ,  $K_I^l$ ,  $K_p^e$ , and  $K_I^e$  are chosen so that the system is robust to various realizations of the rate-utility parameters. The same values of the PI gains are chosen for all programs. A similar analysis can be done when buffer levels are controlled.

## V. EXPERIMENTAL TESTS

### A. Example of application context

A typical application scenario for the proposed rate control system is Mobile TV using the evolved MBMS standard [26]. The MBMS architecture is composed of three main entities: BM-SC, MBMS-GW and MCE. The Multicast/Broadcast Service Center (BM-SC) is a node that serves as entry point for the content providers delivering the video sources, used for service announcements, session management. The MANE, considered in the paper in charge of choosing the encoding and the transmission rates, may be located at the Broadcast/Multicast source at the entrance of the BM-SC node. The MBMS-Gateway (GW) is an entity responsible for distributing the traffic across the different eNBs belonging to the same broadcast area. It ensures that the same content is sent from all the eNBs by using IP Multicast. The Multi-cell/multicast Coordination Entity (MCE) is a logical entity, responsible for allocation of time and frequency resources for multi-cell MBMS transmission. As in [27], we assume that the MBMS-GW periodically notifies the MCE about the resource requirements of video streams so that the resources at eNBs can be re-allocated accordingly. Therefore, the BM-SC should ensure that the encoding rate of the multiplex does not violate the already allocated resources. This is obtained thanks to the proposed rate control scheme.

### B. Simulation environment

To illustrate the properties of the proposed controllers, this section describes a simulation of mobile TV delivery in the previously described context. We consider  $N = 6$  video streams, each of 100 s long, extracted from real TV programs. Interview<sup>3</sup> (Prog 1), Sport<sup>4</sup> (Prog 2), Big Buck Bunny<sup>5</sup> (Prog 3), Nature Documentary<sup>6</sup> (Prog 4), Video Clip<sup>7</sup> (Prog 5), and an extract of *Spiderman*<sup>8</sup> (Prog 6) in 4CIF (704 × 576) format are encoded with x.264 [28] at a frame rate  $F = 30$  fps. GoPs of 10 frames are considered, thus the GoP duration is  $T = 0.33$  s. The videos, already encoded using MPEG-4, have been converted to YUV format using *ffmpeg* [29]. The

<sup>3</sup><http://www.youtube.com/watch?v=l2Y5nIbvHLs>

<sup>4</sup><http://www.youtube.com/watch?v=G63TOHluqno>

<sup>5</sup><http://www.youtube.com/watch?v=YE7VzILtp-4>

<sup>6</sup><http://www.youtube.com/watch?v=NNGDj9IeAuI>

<sup>7</sup><http://www.youtube.com/watch?v=rYEDA3JcQqW>

<sup>8</sup><http://www.youtube.com/watch?v=SYFFVxcRDbQ>

average rate and PSNR of the streams encoded by x.264 with a constant quantization parameter  $QP = 3$  are provided in Table I.

Video	Rate (kbit/s)	PSNR (dB)	Activity
Prog 1	1669.9	46.06	low
Prog 2	4929.1	44.23	high
Prog 3	3654.6	44.56	high
Prog 4	2215.1	44.61	low
Prog 5	2811.4	46.37	medium
Prog 6	3315.9	46.53	high

Table I  
AVERAGE RATE AND PSNR OF THE SIX CONSIDERED VIDEO STREAMS  
(ENCODING WITH X.264 AND CONSTANT  $QP = 3$ ).

The videos are then processed with the proposed control system operating at the GoP level. Initially, all buffers contain three encoded GoPs corresponding to a buffering delay of 1 s. The size  $B_{\max}$  of the buffers is taken large enough to support the variations of its level, occurring, *e.g.*, during scene changes. Here, their size in bits is  $B_{\max} = 4$  Mbits. The reference buffer level in bits is  $B_0 = 400$  kbits and the reference delay is  $\tau_0 = 1.5$  s. This reference delay is consistent with a typical switching time of less than 2 s, as expected in MBMS Television services [30]. The channel rate is  $R^c = 4$  Mbps. The encoding rates are initially considered equal to  $R_0 = R^c/N$ . These rates correspond to the output value of the rate control process provided to each video server. The encoder is then in charge of adjusting its encoding parameters to achieve the target bit rate.

The encoding/transcoding rates are sent to the video encoders which have to choose the encoding parameters for the next VU. In the considered simulation, the video quality is in terms of PSNR or of SSIM of the encoded VUs. This quality metric is transmitted to the MANE in the packet headers. Note that the utility model (3) is only required to characterize the stability of the system and to tune the control parameters. Once the parameters have been chosen off-line, there is no need to know precisely the model (3) within the MANE.

The proposed *quality fair* (QF) video delivery system is compared to a *transmission rate fair* (TRF) controller which provides equal transmission rate to the  $N$  video streams. In the TRF scheme, the encoding rate is controlled to limit the buffer level/delay discrepancy.

A comparison is also performed with a *utility max-min fair* (UMMF) approach [17], with a proportional transmission rate control limiting the buffer level discrepancy. In the UMMF approach, the MANE tries to find the set of encoding rates for the next VU that maximizes the minimum utility. The following constrained optimization problem is then considered

$$\begin{aligned} \mathbf{R}^e(j+1) = & \arg \max_{R_1^e, \dots, R_N^e} \min \{f(\mathbf{a}_1(j), R_1^e), \dots, f(\mathbf{a}_N(j), R_N^e)\} \\ & \text{such that } \sum_{i=1}^N R_i^e = R^c. \end{aligned} \quad (35)$$

Solving (35) requires the availability at the MANE of all rate-utility characteristics (or at least all vectors of parameters  $\mathbf{a}_i(j)$ ) of the previously encoded VUs, contrary to the QF

approach, where only the actual utility of the VUs is needed. The fact that  $\mathbf{a}_i(j)$  is used in (35) for the evaluation of the encoding rate at time  $j+1$  accounts for the possibility for the MANE to get only rate-utility characteristics of previously encoded and already received VUs. Once the value of the encoding rates  $R_1^e, \dots, R_N^e$  are derived, a proportional (P) controller for the transmission rate is applied to evaluate the transmission rate allocated to each video stream

$$R_i^t(j) = R_0 + K_p^t(B_i(j) - B_0), \quad (36)$$

where  $K_p^t$  is the proportional correction gain.

In this section different cases are considered: Both buffer level and buffering delay are addressed separately including stability analysis and results for different utility metrics. Then, the robustness of the proposed control system is analyzed by considering variations of the channel rate as well as of the number of video programs.

### C. Control of the buffer level

We first focus on the system performance when the buffer level (in bits) is used to update the encoding rate.

1) *PSNR-rate model*: To tune the PI controllers of the QF system, the first utility function considered is the PSNR of each GoP. As in [31], a logarithmic PSNR-rate model is used

$$\begin{aligned} U_i(j) = P_i(j) = & f(\mathbf{a}_i(j), R_i^e(j)) \\ = & a_i^{(1)}(j) \log(a_i^{(2)}(j) R_i^e(j)), \end{aligned} \quad (37)$$

with  $P_i(j)$  the PSNR of the GoP at time  $j$  for the  $i$ -th stream. For the  $N = 6$  considered programs, the entries of  $\mathbf{a}_i(j)$  are estimated for each GoP using four encoding trials. An example of the accuracy of the PSNR-rate model (38)

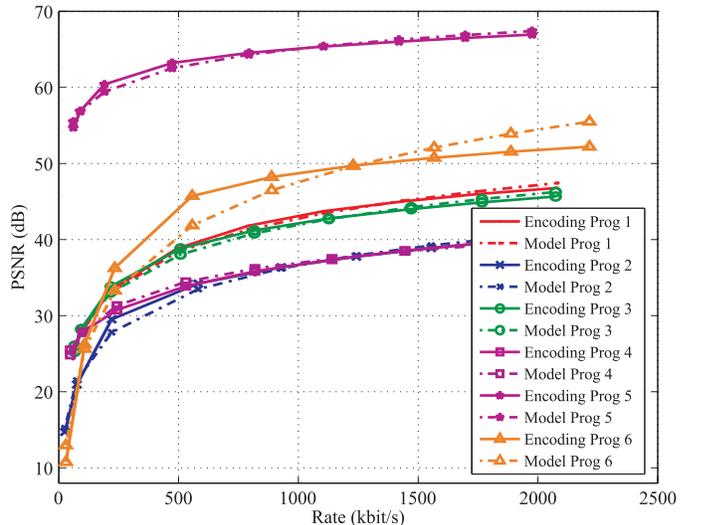


Figure 4. PSNR-rate characteristics and models for the first GoP of the  $N = 6$  considered programs

is shown in Figure 4, when applied on the first GoP of the six considered programs, with parameters estimated from encoding trials performed at 80 kb/s, 200 kb/s, 800 kb/s, and 2 Mb/s. Figure 4 illustrates the ability of (38) to predict the PSNR over a wide range of rates. In addition, the correlation

coefficient  $r^2$  between experimental and predicted PSNR-rate points is evaluated as

$$r^2 = \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} \quad (38)$$

with  $\sigma_x^2 = \sum_{k=1}^n (x_k - \bar{x})^2$ ,  $\sigma_y^2 = \sum_{k=1}^n (y_k - \bar{y})^2$ , and  $\sigma_{xy}^2 = \sum_{k=1}^n (y_k - \bar{y})(x_k - \bar{x})$ , where  $n$  is the number of rates for each program at which the PSNR has been evaluated ( $x_k$ ) and predicted ( $y_k$ ) using (38), and where  $\bar{x}$  and  $\bar{y}$  are the average values of the  $x_k$ 's and of the  $y_k$ 's. For the six programs, for  $n = 7$ , the rate values are 80 kb/s, 130 kb/s, 200 kb/s, 500 kb/s, 800 kb/s, 1.4 Mb/s, and 2 Mb/s, the correlation coefficients are  $r^2 = [0.998, 0.996, 0.997, 0.996, 0.992, 0.985]$  illustrating to good fit by (38) of the PSNR-rate characteristics.

2) *Controller design and stability analysis*: The values of the parameter vector  $\mathbf{a}_i(1)$ ,  $i = 1, \dots, 6$ , obtained for the first GoP of the  $N = 6$  programs are

$$\begin{aligned} \mathbf{a}_1(1) &= \begin{pmatrix} 1.11 \\ 0.15 \end{pmatrix}, \mathbf{a}_2(1) = \begin{pmatrix} 1.90 \\ 0.17 \end{pmatrix}, \mathbf{a}_3(1) = \begin{pmatrix} 0.76 \\ 0.17 \end{pmatrix}, \\ \mathbf{a}_4(1) &= \begin{pmatrix} 0.09 \\ 0.24 \end{pmatrix}, \mathbf{a}_5(1) = \begin{pmatrix} 2.50 \\ 0.17 \end{pmatrix}, \mathbf{a}_6(1) = \begin{pmatrix} 0.07 \\ 0.20 \end{pmatrix}. \end{aligned}$$

Once  $f$  is specified, one may characterize the system equilibrium. The vector of rates at equilibrium  $(R_1^{\text{e,eq}}, \dots, R_N^{\text{e,eq}})^T$  is obtained by solving the system of equations in the state-space representation at equilibrium.  $\Xi$  and  $\Gamma$  are derived from (29), (30) and (38) as follows

$$\Xi = \begin{pmatrix} \log(a_{1,2}(1) R_1^{\text{e,eq}}) & \frac{a_{1,1}(1)}{a_{1,2}(1)} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \log(a_{2,2}(1) R_2^{\text{e,eq}}) & \frac{a_{2,1}(1)}{a_{2,2}(1)} & 0 & \dots & 0 \\ \vdots & \vdots & 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & \log(a_{N,2}(1) R_N^{\text{e,eq}}) & \frac{a_{N,1}(1)}{a_{N,2}(1)} \end{pmatrix} \quad (39)$$

and

$$\Gamma = \text{diag} \left( \frac{a_{1,1}(1)}{R_1^{\text{e,eq}}(1)}, \dots, \frac{a_{N,1}(1)}{R_N^{\text{e,eq}}(1)} \right). \quad (40)$$

The gains of the PI controllers have to be chosen so that the roots of

$$d(z) = \det(z\mathbf{I} - \mathbf{A}) \quad (41)$$

remains within the unit circle, for various rate-utility characteristics of the VUs. In (41),  $\mathbf{A}$  may correspond to  $\mathbf{A}^b$ , the linearized state matrix when considering buffer level control, or to  $\mathbf{A}^\tau$  with buffering delay control.

To increase the robustness of the proposed approach to variations of the rate-utility characteristics,  $K = 10$  realizations of  $N = 4$  random parameter vectors of the PSNR-rate model obtained as follows

$$\begin{aligned} a_{i,1}^{(k)} &= \frac{1}{N} \sum_{i=1}^N a_{i,1}(1) + \eta_{i,1}^{(k)}, \\ a_{i,2}^{(k)} &= \frac{1}{N} \sum_{i=1}^N a_{i,2}(1) + \eta_{i,2}^{(k)}, \end{aligned} \quad (42)$$

for  $k = 1, \dots, K$ . In (43),  $\eta_{i,1}^{(k)}$  and  $\eta_{i,2}^{(k)}$  are realizations of zero-mean Gaussian variables with variance  $\sigma_1^2 = 6.25 \times 10^{-2}$  and  $\sigma_2^2 = 2.25 \times 10^{-4}$ . The resulting PSNR-rate characteristics obtained using (43) are represented in Figure 5. These random realizations describe quite well the variability of actual PSNR-rate characteristics represented in Figure 4.

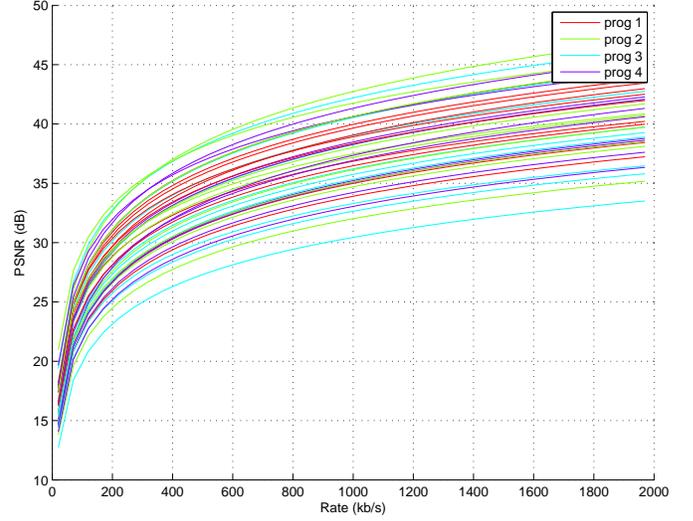


Figure 5. Superposition of the  $K = 10$  realizations of the  $N = 4$  random PSNR-rate characteristics

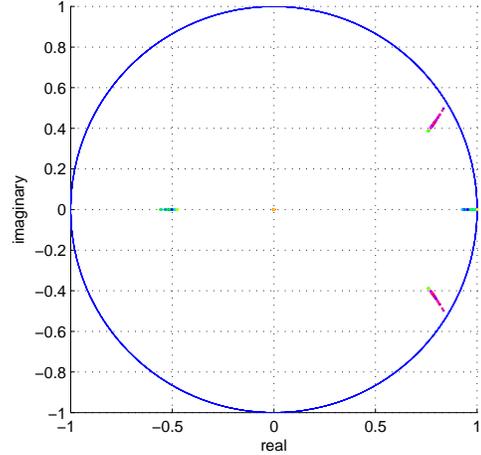


Figure 6. Location of the roots of  $\det(z\mathbf{I} - \mathbf{A}^b)$  for the  $K = 10$  realizations of the  $N = 4$  random PSNR-rate characteristics

A random search of the control parameters is then performed. Among the values providing stability for the  $K$  random PSNR-rate characteristics, the one with the roots farthest away from the unit circle is selected to provide good transients.

The tuning is performed for  $N = 4$ . Good transient behaviors have been obtained with  $K_P^{\text{e,b}} = 666$ ,  $K_I^{\text{e,b}} = 33$ ,  $K_P^{\text{t}} = 66 \times 10^3$ , and  $K_I^{\text{t}} = 1300$ . The position of the roots corresponding to the video characteristic represented in Figure 5 are in Figure 6.

Figure 6 shows that all roots remain within the unit circle. This result does not prove the robustness of the proposed choice of the control parameters, but shows that this choice leads to a system reasonably robust to changes of the characteristics of the transmitted programs. Nevertheless, some of the roots are located quite near the stability limit, which will lead to quite long transients.

3) *Simulation results:* The x.264 video encoder is used to perform on-line compression of the different programs with the rate targets provided by the encoding rate controllers. Within the video coder, a two-pass rate control is performed to better fit the target encoding rate. The target and obtained encoding rates may however be slightly different. The system performance is first measured in terms of average buffer level discrepancy  $\Delta_B$  (in bits) with respect to  $B_0$ , variance of the buffer level  $\sigma_B^2$  (in bits<sup>2</sup>), PSNR discrepancy  $\Delta_P$  (in dB), and PSNR variance  $\sigma_P^2$  (in dB<sup>2</sup>), with

$$\begin{aligned} \Delta_B &= \frac{1}{NM} \sum_{n=1}^N \sum_{l=1}^M (B_n(l) - B_0), \\ \sigma_B^2 &= \frac{1}{N} \sum_{n=1}^N \sum_{l=1}^M (B_n(l) - B_0 - \Delta_B)^2, \\ \Delta_{PSNR} &= \frac{1}{NM} \sum_{n=1}^N \sum_{l=1}^M (P_n(l) - \bar{P}(l)), \\ \sigma_{PSNR}^2 &= \frac{1}{N} \sum_{n=1}^N \sum_{l=1}^M (P_n(l) - \bar{P}(l) - \Delta_{PSNR})^2, \end{aligned} \quad (43)$$

where  $\bar{P}(l) = \frac{1}{N} \sum_{n=1}^N P_n(l)$  and  $M$  is the number of GoPs in the video streams.

The results with  $N = 6$  and  $R^c = 4$  Mbit/s are summarized in Table II in the three cases: TRF, UMMF, and QF where  $\Delta_B$  are in kbits and  $\sigma_B^2$  in kbit<sup>2</sup>. The PI controllers used for the transmission rate control loop reduce the PSNR discrepancy between the programs at a price of some increase of the buffer level discrepancy and variance.

	$K_p^{c,b}, K_I^{c,b}$	$K_p^t, K_I^t$	$ \Delta_B $	$\sigma_B^2$	$ \Delta_P $	$\sigma_P^2$
TRF	666, 0	0, 0	31.5	2.1	3.1	9.8
UMMF	0, 0	3, 0	75	2	2.7	13.9
QF	666, 33	$(66, 1.3)10^3$	53.07	6.1	1.5	6.7

Table II  
PERFORMANCE WHEN USING TRF AND QF CONTROLLERS WHEN CONTROLLING THE BUFFER LEVELS FOR  $N = 6$ .

Figure 7 represents the evolution of the PSNR of the  $N$  programs over 300 GoPs using the TRF (left) and the QF (right) controllers when  $N = 2$ ,  $N = 4$ , and  $N = 6$  with a constant channel rate  $R^c = 4$  Mbit/s. The same controller parameters  $K_p^{c,b} = 666$ ,  $K_I^{c,b} = 33$ ,  $K_p^t = 66 \times 10^3$ , and  $K_I^t = 1300$  are used for  $N = 2$ ,  $N = 4$ , and  $N = 6$ .

When  $N = 2$ , the average PSNR of Prog 2, characterized by high activity level, is improved from 36 dB to 41 dB, leading to a significant improvement of the video quality. This is at the price of PSNR degradation of Prog 1, characterized by low activity level, from 45 dB to 41 dB. This still corresponds to a very good quality. PSNR fairness improvements are also obtained when  $N = 4$  and  $N = 6$ .

Figure 8 shows the evolution of the buffer level of the  $N$  programs over 300 GoPs using the TRF (left) and the QF (right) controllers for  $N = 2$ ,  $N = 4$ , and  $N = 6$ . The discrepancy between the buffer level and the reference level  $B_0$  remains limited for most of the time. When only the encoding rate is controlled, corresponding to the TRF controller, the buffer level stabilizes around  $B_0$ . The buffer level variations increase using the QF controller due to the interactions of the encoding rate and transmission rate control loops.

Figure 9 shows the evolution of the buffer level (left) and of the PSNR (right) of the  $N$  programs over 300 GoPs when the UMMF technique is used. The choice of the proportional gain

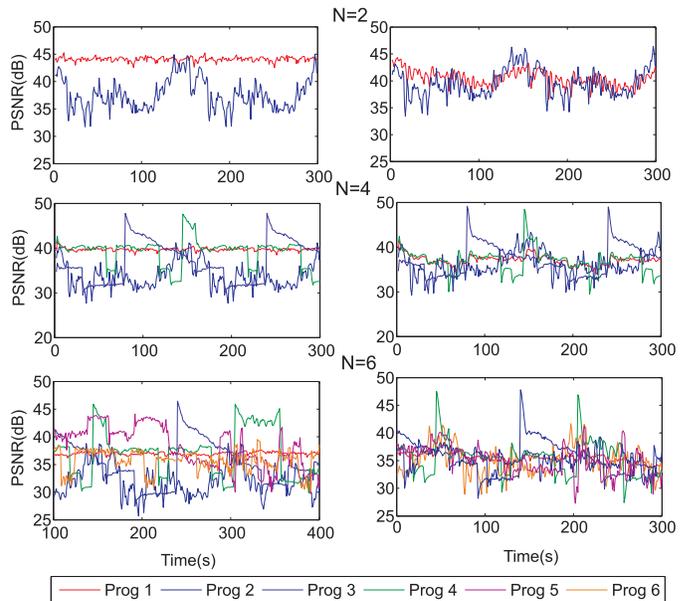


Figure 7. Evolution of the PSNR for  $N = 2$ ,  $N = 4$ , and  $N = 6$  using TRF (left) and QF (right) controllers when controlling the buffer levels

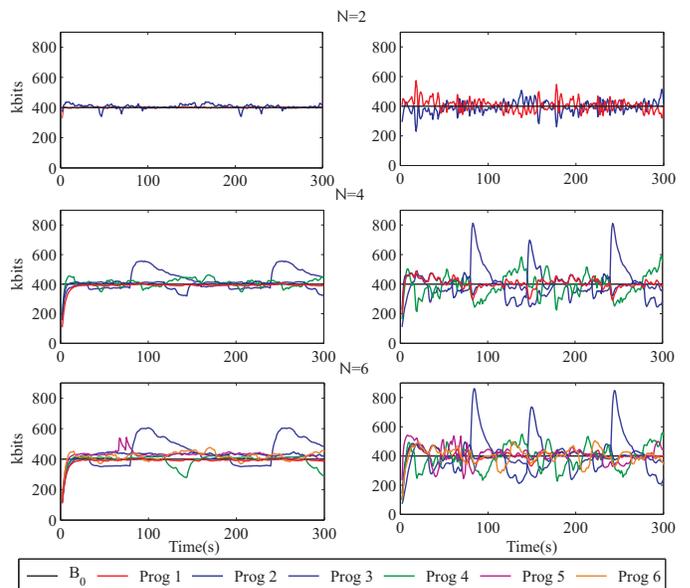


Figure 8. Evolution of the buffer level for  $N = 2$ ,  $N = 4$ , and  $N = 6$  using TRF (left) and QF (right) controllers when controlling the buffer levels

for the transmission rate controller has no significant impact on the quality fairness, provided that the buffers remain full. A reference buffer level of 400 kbits has been used to allow a satisfying behavior of the transmission rate control loop. The PSNR remains around 40 dB, but the average variance is of the same order of magnitude as that of the TRF solution, see Table II. This mainly comes from the target encoding rate evaluation on (one step) outdated PSNR-rate characteristics.

4) *SSIM-rate model:* The quality fairness is also addressed considering the SSIM metric. To tune the PI controllers, an

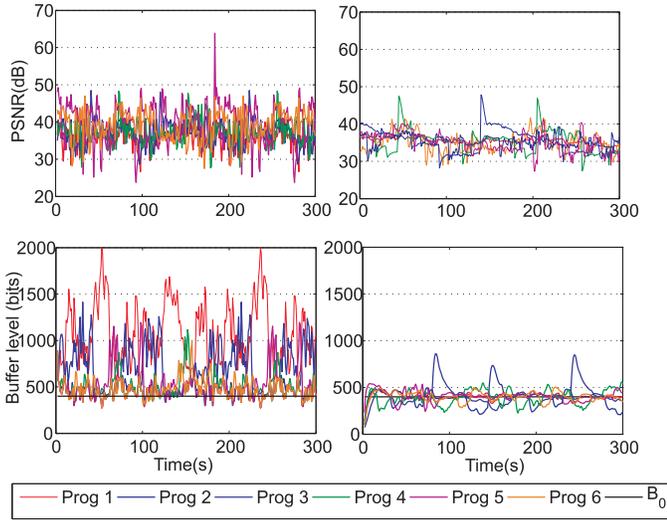


Figure 9. Evolution of the buffer level and of the PSNR (right) for  $N = 6$  using the UMMF technique (left) and using the QF controllers (right) with a transmission rate control using the buffer levels

arctan SSIM-Rate utility model is considered

$$\begin{aligned} U_i(j) &= S_i(j) = f(\mathbf{a}_i(j), R_i^e(j)) \\ &= a_i^{(1)}(j) \operatorname{atan}(a_i^{(2)}(j) R_i^e(j)) \end{aligned} \quad (44)$$

where  $S_i(j)$  is the SSIM of program  $i$  at time  $j$ . As before, the two entries of each  $\mathbf{a}_i(j)$  are derived from four encoding trials performed on each of the  $N$  considered programs. The resulting values of the parameters are

$$\begin{aligned} \mathbf{a}_1(1) &= \begin{pmatrix} 0.64 \\ 0.037 \end{pmatrix}, \mathbf{a}_2(1) = \begin{pmatrix} 0.61 \\ 0.029 \end{pmatrix}, \mathbf{a}_3(1) = \begin{pmatrix} 0.64 \\ 0.034 \end{pmatrix}, \\ \mathbf{a}_4(1) &= \begin{pmatrix} 0.62 \\ 0.017 \end{pmatrix}, \mathbf{a}_5(1) = \begin{pmatrix} 0.64 \\ 0.22 \end{pmatrix}, \mathbf{a}_6(1) = \begin{pmatrix} 0.64 \\ 0.044 \end{pmatrix}. \end{aligned}$$

Figure 10 compares the actual SSIM-Rate characteristics and those obtained using the model (45) for the first GoP of the six considered programs. The model (45) is able to predict accurately the SSIM over a large range of rates. The correlation coefficient for the six programs is  $r^2 = [0.99, 0.97, 0.99, 0.98, 0.99, 0.99]$  using the same rate values in the PSNR-rate model estimation, which confirms the accuracy of the SSIM-rate model.

Using the SSIM-rate utility function (45), one is able to get the vector  $(R_1^{e,\text{eq}}, \dots, R_N^{e,\text{eq}})^T$  of encoding rates as the solution of (24).  $\Xi$  and  $\Gamma$  are derived from (29), (30) and (45) as follow

$$\Xi = \begin{pmatrix} \operatorname{atan}(a_1^{(2)} R_1^{e,\text{eq}}) & \frac{a_1^{(1)} R_1^{e,\text{eq}}}{1+(a_1^{(2)} R_1^{e,\text{eq}})^2} & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & \dots & 0 & \operatorname{atan}(a_N^{(2)} R_N^{e,\text{eq}}) & \frac{a_N^{(1)} R_N^{e,\text{eq}}}{1+(a_N^{(2)} R_N^{e,\text{eq}})^2} \end{pmatrix} \quad (45)$$

and

$$\Gamma = \operatorname{diag} \left( \frac{a_1^{(1)} a_1^{(2)}}{1+(a_1^{(2)} R_1^{e,\text{eq}})^2}, \dots, \frac{a_N^{(1)} a_N^{(2)}}{1+(a_N^{(2)} R_N^{e,\text{eq}})^2} \right) \quad (46)$$

The choice of the parameters of the controllers is done as in Section V-C2. Good transient behaviors have been obtained with  $K_P^{c,b} = 666$ ,  $K_I^{c,b} = 33$ ,  $K_P^l = 66 \times 10^4$ , and  $K_I^l =$

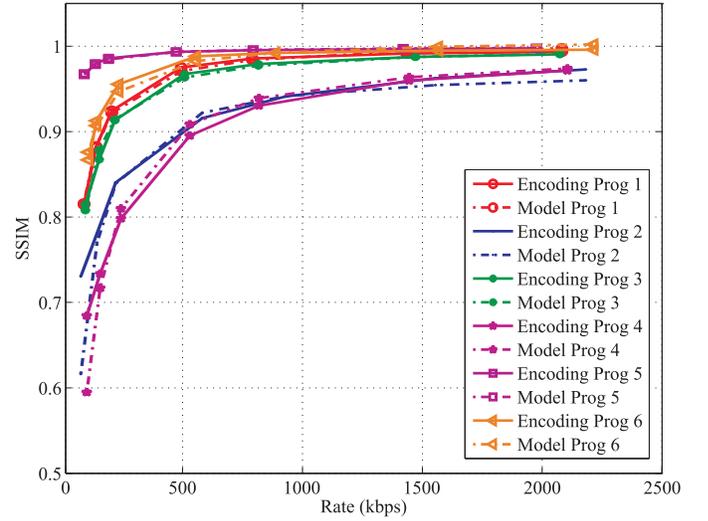


Figure 10. SSIM-Rate model for the six considered programs

$1.3 \times 10^4$ . For this choice of the control gains, Figure 11 shows the minimum, average, and maximum values of the PSNR (left) and SSIM (right) over all GoPs of the  $N = 6$  programs. The QF controller improves quality fairness especially for the most demanding videos, such as Prog 2. The proposed QF controller improves also the minimum achieved quality for these programs. In fact, even if events corresponding to these minimum quality happens only few times, the user perception is sometime dominated by the worst experience, rather than the average. The price to be paid is, as expected, a decrease of the quality of the less demanding programs.

#### D. Control of the buffering delays

This part focuses on the system performance when the buffering delays are used to evaluate the encoding rates.

1) *Controller design*: The same PSNR-rate utility model as in (38) is considered in this section. Thus the matrices  $\Xi$  and  $\Gamma$  are those in (39) and (40).

The choice of the parameters of the two PI controllers is again done as in Section V-C2. Now, the roots of  $\det(z\mathbf{I} - \mathbf{A}^\tau)$  have to remain within the unit circle. Good transient behaviors have been obtained for  $N = 2$ ,  $N = 4$ , and  $N = 6$  with  $K_P^l = 66 \times 10^3$ ,  $K_I^l = 2600$ ,  $K_P^{e\tau} = 66 \times 10^3$ , and  $K_I^{e\tau} = 1300$ .

In parallel, the parameter  $\alpha$  in (9) is tuned to provide the best estimate of the buffering delay. Figure 12 represents the means square error  $\operatorname{MSE}(\tilde{\tau}, \tau)$  between the actual buffering delay  $\tau$  and the estimated one  $\tilde{\tau}$  as a function of  $\alpha$ . The value  $\alpha = 0.2$  provides the best estimate. The evolution with time of the actual buffering delay  $\tau$  and of its estimate  $\tilde{\tau}$  is represented in Figure 13 for  $N = 4$  using  $\alpha = 0.2$  and the QF controller for the PSNR fairness. For this choice of  $\alpha$ , the estimate provided by (9) for the four video sequences is quite good for most of the time.

2) *Results*: The performance of the QF controller using  $N = 2$ ,  $N = 4$ , and  $N = 6$  programs is evaluated in all cases with a transmission rate  $R^c = 4$  Mps.

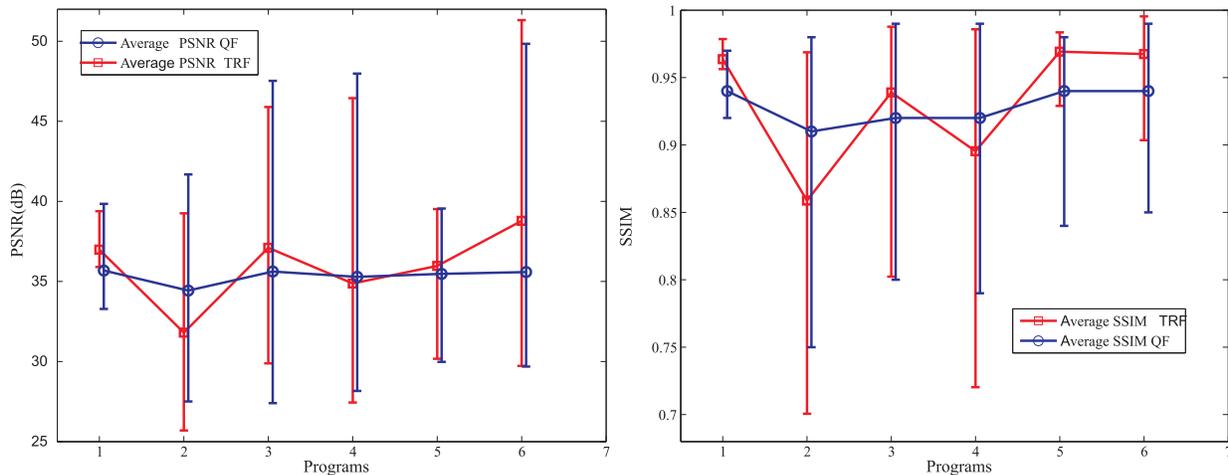


Figure 11. Minimum, average, and maximum values of the PSNR (left) and SSIM (right) over all GoPs for the TRF and the QF controllers using  $N = 6$  programs when controlling the buffer levels

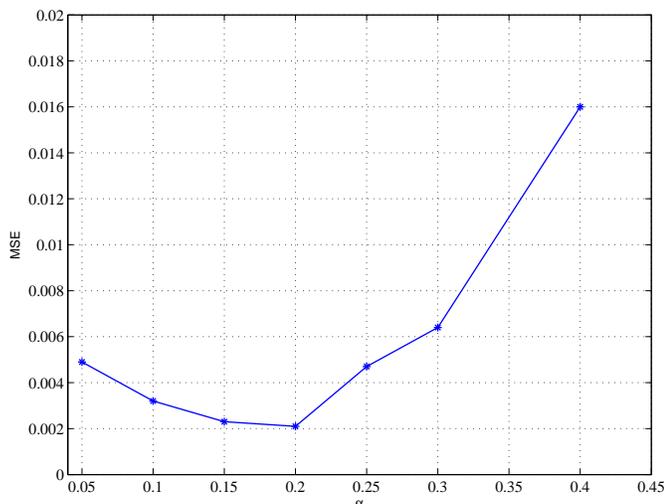


Figure 12. MSE of the estimated buffering delay as a function of  $\alpha$ .

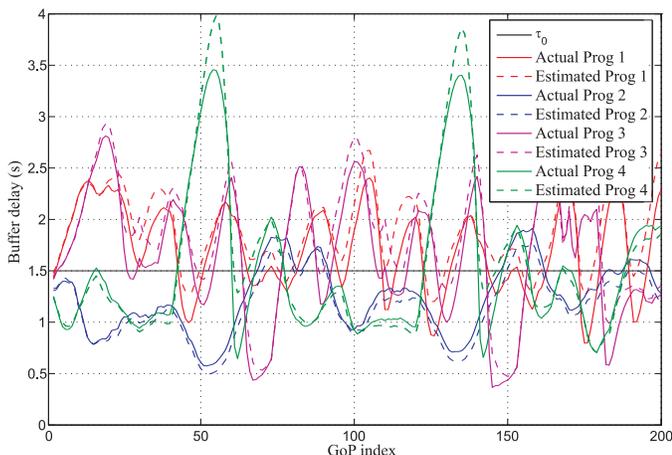


Figure 13. Evolution of the actual buffering delay  $\tau$  and of its estimate  $\tilde{\tau}$  for  $N = 4$  using  $\alpha = 0.2$ .

We evaluate the PSNR discrepancy  $\Delta_P$  (in dB) and the

PSNR variance  $\sigma_P^2$  (in  $\text{dB}^2$ ) defined in (43). Additional performance measures are the average delay discrepancy

$$\Delta_\tau = \frac{1}{N} \sum_{n=1}^N \left( \frac{1}{M} \sum_{l=1}^M (\tau_n(l) - \tau_0) \right) \quad (47)$$

of the buffering delay with respect to  $\tau_0$  and the variance of the buffering delay

$$\sigma_\tau^2 = \frac{1}{N} \sum_{n=1}^N \left( \frac{1}{M} \sum_{l=1}^M (\tau_n(l) - \tau_0 - \Delta_\tau)^2 \right), \quad (48)$$

where  $M$  is the number of GoPs in the video streams. Results with  $N = 6$  are summarized in Table III in the two cases: QF and TRF controllers. Here again, one notices that using PI

	$K_p^{e,\tau}, K_l^{e,\tau}$	$K_p^l, K_l^l$	$\Delta_\tau$	$\sigma_\tau^2$	$\Delta_P$	$\sigma_P^2$
TRF	$66 \times 10^3, 0$	0, 0	0.25	0.12	3.8	10.5
QF	$66 \times 10^3, 1300$	$66 \times 10^3, 2600$	0.6	0.35	2	10

Table III  
PERFORMANCE OF QF AND TRF CONTROLLERS WHEN CONTROLLING THE BUFFERING DELAYS FOR  $N = 6$ .

controllers for the transmission rate control loop reduces the PSNR discrepancy between the programs at the price of some increase of the buffering delay discrepancy and variance.

Figure 14 represents the evolution of the PSNR when considering the TRF controller (left) and the proposed QF controller (right)  $N = 2$ ,  $N = 4$ , and  $N = 6$  programs. The proposed QF controller reduces the PSNR discrepancy between the  $N$  programs compared to the TRF controller. Compared to Figure 7, the control with the buffer level appears to be less reactive. For example, when  $N = 2$ , to improve the PSNR of the second program, the PSNR of the first program has to be decreased. In Figure 7, the PSNRs are almost immediately adjusted. This is done with some delay in Figure 14. This may be due to the difficulty to accurately estimate the buffering delay. A better response could be obtained by increasing  $K_p^{e,\tau}$ , which relates the buffering delay and the encoding rate. This, however, would be at the price of

a loss in robustness of the global system to variations of the PSNR-rate characteristics of the programs.

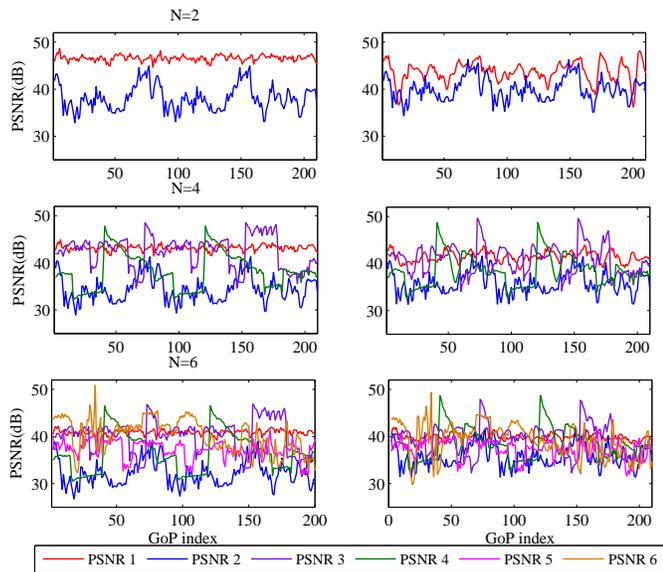


Figure 14. Evolution of the PSNR for  $N = 2$ ,  $N = 4$ , and  $N = 6$  using TRF (left) and QF (right) controllers when controlling the buffer delays.

Figure 15 represents the evolution of the buffering delays of  $N$  programs of 300 GoPs using the TRF (left) and the QF (right) controllers for  $N = 2$ ,  $N = 4$ , and  $N = 6$ . With the TRF controller, the buffering delays reach rapidly  $\tau_0$  and show a reduced variance compared to a system with a QF controller. The larger variations of the buffering delay for the QF controller are due to the interactions of both control loops (encoding rate and transmission rate). Again,  $K_p^{e,\tau}$  appears to be too low: large deviations of the buffering delay are required to reach PSNR fairness.

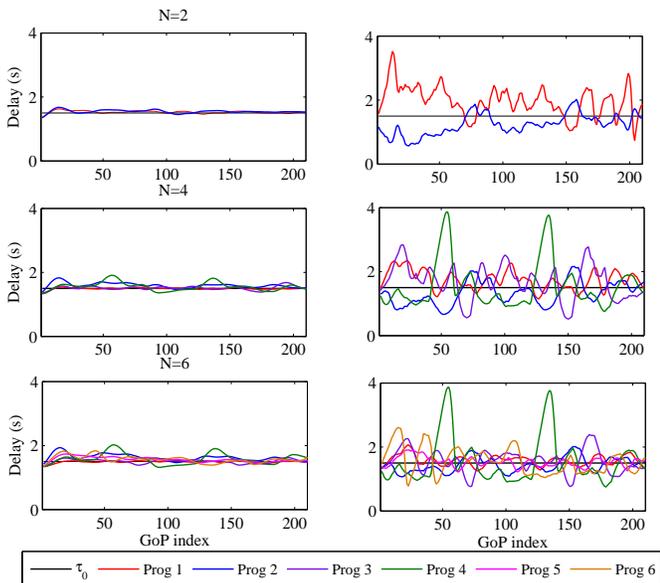


Figure 15. Evolution of the buffering delay for  $N = 2$ ,  $N = 4$ , and  $N = 6$  using TRF (left) and QF (right) controllers when controlling the buffer delays.

*E. Robustness of the proposed solution to variations of the number of users and of the channel rate*

In this section, the robustness of the proposed control system (buffering delay control) is evaluated with respect to variations of the channel rate and of the number of transmitted video programs. Similar results are obtained when buffer levels are controlled.

First, the number  $N$  of transmitted video programs evolves with time (left). Second, the rate of the channel switches between  $R^c = 3.5$  Mbits/s and  $R^c = 5$  Mbits/s (right), see Figure 16. The PSNR is used as quality measure. When a new video program is transmitted, initially, it has no transmission rate allocated by the MANE (since at time  $j$  the controller derives the encoding rate for time  $j + 1$ ). Thus, we choose to set the encoding rate at that time as  $R^c/N$ . In Figure 16 (left), Prog 4 is not transmitted between GoP 35 and 65. The same values for the gains of the PI controllers are used here as in Section V-D.

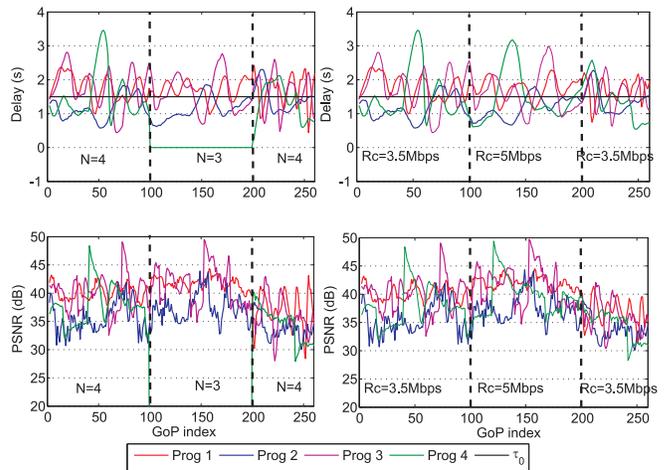


Figure 16. System performance using PI controllers while multiplexing four video programs using the proposed QF controller when considering variations of the channel rate (left) and of the number of programs (right).

When the channel rate increases or when a video program is no more transmitted, the bandwidth allocation adapts rapidly to this change by providing more rate to programs with low video quality (here Prog 2). When the channel rate decreases or when a new video program is transmitted, the bandwidth allocation performs well, showing the robustness of the proposed control system to variations of the channel rate and to the number of transmitted video programs.

## VI. CONCLUSIONS AND PERSPECTIVES

In this paper, we propose encoding and transmission rate controllers for the transmission of several video streams targeting similar video quality between streams as well as efficient control of the buffering delay. The controlled system is modeled with a discrete-time non-linear state-space representation. PI controllers for the transmission rate and the encoding rate control are considered. The delay introduced by the network propagation between the MANE and the encoders is taken into account. This allows to test the stability of the control system

in presence of feedback delay. Simulation results show that the quality fairness (measured with PSNR or SSIM) is improved compared to a solution providing an equal transmission rate allocation. Moreover, the jitter of the buffering delay remains reasonable. The robustness to variations of the characteristics of the channel and of the number of transmitted programs has been shown experimentally.

Simulations are performed at the GoP granularity. Control at the frame level should be considered, however this may require a better consideration of the communication delay between the MANE and the encoders which may be variable with the time. This would also require to better account for the delay, which significantly impedes the behavior of the global control system, especially when controlling the buffering delay. Tools devoted to the control of time-delay systems may be useful in this context, see, e.g.,[32].

## REFERENCES

- [1] "Cisco visual networking index: Forecast and methodology, 2008-2013," Tech. Rep., Cisco, 2009.
- [2] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600 – 612, April 2004.
- [3] Z.L. Zhang, S. Nelakuditi, R. Aggarwal, and R.P. Tsang, "Efficient selective frame discard algorithms for stored video delivery across resource constrained networks," *Real-Time Imaging*, vol. 7, no. 3, pp. 255 – 273, March 2001.
- [4] T-L. Lin, Y. Zhi, S. Kanumuri, P.C. Cosman, and A.R. Reibman, "Perceptual quality based packet dropping for generalized video gop structures," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2009, pp. 781 – 784.
- [5] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1533 – 1544, December 2005.
- [6] Z. He and D. O. Wu, "Linear rate control and optimum statistical multiplexing for H.264 video broadcast," *IEEE Transactions on Multimedia*, vol. 10, no. 7, pp. 1237 – 1249, November 2008.
- [7] S. K. Srinivasan, J. Vahabzadeh-Hagh, and M. Reisslein, "The effects of priority levels and buffering on the statistical multiplexing of single-layer H.264/AVC and SVC encoded video streams," *IEEE Transactions on Broadcasting*, vol. 56, no. 3, pp. 281 – 287, September 2010.
- [8] E. Maani and A.K. Katsaggelos, "Unequal error protection for robust streaming of scalable video over packet lossy networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 3, pp. 407 – 416, March 2010.
- [9] Y. Li, Z. Li, M. Chiang, and A.R. Calderbank, "Content-aware distortion-fair video streaming in congested networks," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1182 – 1193, October 2009.
- [10] N. Changuel, B. Sayadi, and M. Kieffer, "Joint encoder and buffer control for statistical multiplexing of multimedia contents," in *IEEE Globecom*, December 2010, pp. 1 – 6.
- [11] C.W. Wong, O.C. Au, and H.K. Lam, "PID-based real-time rate control," in *IEEE International Conference on Multimedia and Expo*, June 2004, vol. 1, pp. 221 – 224.
- [12] Z. Guo and R. Rojas-Cessa, "Analysis of a flow control system for a combined input-crosspoint buffered packet switch," in *Workshop on High Performance Switching and Routing*, May 2005, pp. 336 – 340.
- [13] Y-H. Zhang, K.H. Li, C.Q. Xu, and L.M. Sun, "Joint rate allocation and buffer management for robust transmission of VBR video," *Acta Automatica Sinica*, vol. 34, no. 3, pp. 1 – 7, March 2008.
- [14] Y. Huang, S. Mao, and S.F. Midkiff, "A control-theoretic approach to rate control for streaming videos," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1072 – 1081, October 2009.
- [15] Y. Huang and S. Mao, "Analysis and design of a proportional-integral rate controller for streaming videos," in *IEEE Globecom*, November 2009, pp. 5530 – 5535.
- [16] Y. Zhou, Y. Sun, Z. Feng, and S. Sun, "PID-based bit allocation strategy for h.264/avc rate control," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 3, pp. 184 – 188, March 2011.
- [17] J.W. Cho and S. Chong, "Utility max-min flow control using slope-restricted utility functions," in *IEEE Globecom*, December 2005, vol. 2, pp. 818 – 824.
- [18] S. Cicalo and V. Tralli, "Cross-layer algorithms for distortion-fair scalable video delivery over OFDMA wireless systems," in *IEEE Globecom Workshops*, 2012, pp. 1287 – 1292.
- [19] E. Witrant, C. Canudas-de Wit, D. Georges, and M. Alamir, "Remote stabilization via communication networks with a distributed control law," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1480 – 1485, August 2007.
- [20] Motorola, "Statistical multiplexing over IP StatmuxIP solution," Tech. Rep., Motorola, 2009, XP002565902.
- [21] Harmonic, "IP-based distributed statistical multiplexing solution," Tech. Rep., 2011.
- [22] N. Changuel, B. Sayadi, and M. Kieffer, "Control of distributed servers for quality-fair delivery of multiple video streams," in *ACM international conference on Multimedia*, October 2012, pp. 269 – 278.
- [23] K. Seshadrinathan, R. Soundararajan, A.C. Bovik, and L.K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1427 – 1441, June 2010.
- [24] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G.J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 688 – 703, July 2003.
- [25] K.-B. Kim and H.-J. Kim, "Back-pressure buffering scheme to improve the cell loss property on the output buffered atm switch," in *IEEE Conference on Local Computer Networks*, October 1996, pp. 242 – 248.
- [26] ETSI, "Multimedia Broadcast/Multicast service. (MBMS); UTRAN/GERAN requirements," Tech. Rep., 3GPP TR 25.992-140, 2005.
- [27] V. Vukadinovic and J. Huschke, "Statistical multiplexing gains of H.264/AVC video in E-MBMS," in *International Symposium on Wireless Pervasive Computing*, May 2008, pp. 468 – 474.
- [28] x264 Home Page, "Videolan organization. retrieved 2011-03-11.," 2011.
- [29] ffmpeg Documentation, "http://ffmpeg.org/ffmpeg.html," .
- [30] Association of Radio Industries and Businesses, "Multimedia Broadcast/Multicast Service (MBMS) user services stage 1 (release 8)," Tech. Rep., 3GPP TS, 2008.
- [31] X. Zhu, P. Agrawal, J. Pal Singh, T. Alpcan, and B. Girod, "Rate allocation for multi-user video streaming over heterogenous access networks," in *ACM Multimedia*, Augsburg, September 2007, pp. 37 – 46.
- [32] W. Michiels and S.L. Niculescu, *Stability and Stabilization of Time-Delay Systems: An Eigenvalue-Based Approach*, Society for Industrial & Applied Mathematics, U.S., 2007.



**Nesrine Changuel** (IEEE S'09 - M'12) received her B.S. degree in 2006 and Eng. degree and M.S. degree in electrical engineering in 2008 from the Ecole Nationale Supérieure d'Electronique et de Radioélectricité de Grenoble, France. She obtained a PhD degree in Control and Signal Processing in 2011 from the Paris-Sud University, Orsay. She is currently working as a research engineer in Alcatel Lucent Bell Labs. Her current research interests are in the areas of video coding, Statistical multiplexing of video programs, scalable video, Rate and Distortion model, resource allocation and scheduling, Markov decision processes (MDPs), and reinforcement learning. She is a Member of the IEEE.



**Bessem Sayadi** is a member of technical staff in Multimedia Technology domain at Alcatel-Lucent Bell Labs, France. He is member of Alcatel-Lucent Technical Academy, since 2008. He received M.Sc. (00) and Ph.D. (03) degrees in Control and Signal processing from Supélec, Paris-Sud University, with highest distinction. He worked previously as a postdoctoral fellow in the National Centre for Scientific Research (CNRS), and as a senior researcher engineer in Orange Labs. His main research interests are in the area of broadcast technology (DVB, 3GPP), video coding and transport, and resource allocation algorithms for communication networks. He has authored over 55 publications in journal and conference proceedings and serves as a regular reviewer for several technical journals and conferences. He holds nine patents and has more than twenty patent applications pending in the area of video coding and wireless communications.



**Michel Kieffer** (IEEE M'02 - SM'07) received in 1995 the Agrégation in Applied Physics at the Ecole Normale Supérieure de Cachan. He obtained a PhD degree in Control and Signal Processing in 1999, and the Habilitation à Diriger des Recherches degree in 2005, both from the Paris-Sud University, Orsay.

Michel Kieffer is a full professor in signal processing for communications at the Paris-Sud University and a researcher at the Laboratoire des Signaux et Systèmes, Gif-sur-Yvette. Since 2009, he is also invited professor at the Laboratoire Traitement et

Communication de l'Information, Télécom ParisTech, Paris.

His research interests are in signal processing for multimedia, communications, and networking, distributed source coding, network coding, joint source-channel coding and decoding techniques, joint source-network coding. Applications are mainly in the reliable delivery of multimedia contents over wireless channels. He is also interested in guaranteed and robust parameter and state bounding for systems described by non-linear models in a bounded-error context.

Michel Kieffer is co-author of more than 120 contributions in journals, conference proceedings, or books. He is one of the co-author of the book *Applied Interval Analysis* published by Springer-Verlag in 2001 and of the book *Joint source-channel decoding: A crosslayer perspective with applications in video broadcasting* published by Academic Press in 2009. He is associate editor of *Signal Processing* since 2008, of the *Journal of Communication and Information Systems* since 2011 and of the *IEEE Transactions on Communications* since 2012. In 2011, Michel Kieffer became junior member of the Institut Universitaire de France.