Avoiding Interruptions - a QoE Reliability Function for Streaming Media Applications†

Ali ParandehGheibi, Muriel Médard, Asuman Ozdaglar, Srinivas Shakkottai

***Abstract*—We take an analytical approach to study fundamental rate-delay-reliability trade-offs in the context of media streaming. We consider the probability of interruption in media playback (buffer underflow) as well as the number of initially buffered packets (initial waiting time) as the Quality of user Experience (QoE) metrics. We characterize the optimal trade-off between these metrics as a function of system parameters such as the packet arrival rate and file size, for different channel models. In the first model, we assume packets arrive according to independent Poisson processes from multiple servers or peers. We use random linear network coding to simplify the packet requests at the network layer and avoid duplicate packet reception. This allows us to model the receiver's buffer as a queue with Poisson arrivals and deterministic departures. For this model, we show that for arrival rates slightly larger than the play rate, the minimum initial buffering required to achieve certain level of interruption probability remains bounded as the file size grows. This is not the case when the arrival rate and the play rate match. In the second model, we consider channels with memory, which can be modeled using Markovian arrival processes. We characterize the optimal trade-off curves for the infinite file size case, in such Markovian environments.**

I. Introduction

Peer-to-peer networks (P2P) are a fast-growing means of video delivery. It has been estimated that between 35-90% of Internet bandwidth is consumed by P2P applications [2], [3]. Today, P2P file-sharing networks are seeing a drop in popularity [4], but the original file sharing ideas are being used for video streaming in networks such as PPLive [5] and QQLive [6]. As smart phones become the medium of choice for Internet media access, P2P video distribution over the wireless medium is likely to gain significance.

Our goal in this paper is to understand the fundamental trade-offs among end-user rate, delay and reliability metrics in the context of video streaming. For communication over a noisy channel (physical layer), the following equation captures the essence of such trade-offs from an information theoretic point of view:

$$\text{Probability of error} = e^{-E(R)\cdot(\text{block length})}, \qquad (1)$$

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

P2P video streaming is generally accomplished by dividing the video file into *blocks,* which are then further divided into packets for transmission. After each block is received, it can be played out by the receiver. In order to ensure smooth sequential playback, a fresh block must be received before the current block has been played. If such a fresh block is not available the playback stops, causing a negative user experience. Blocks may be buffered in advance of playing out in order to provide a level of protection against a playback interruption. Hence, there is a trade-off between the initial waiting time and likelihood of playback interruptions.

In this paper, our main objective is to characterize the amount of buffering needed for a target probability of playback interruption over the duration of the video. However, since the wireless channel is unreliable, packets cannot be obtained deterministically. Thus, our question is *how much should we buffer prior to playback in order to account for wireless channel variations?*

We first consider the problem of streaming a media file of finite size from multiple servers (peers) to a single receiver. We assume that each server can effectively transmit packets according to an independent Poisson process. We use random linear network coding to simplify the packet requests at the network layer and avoid duplicate packet reception. This allows us to model the receiver's buffer as an M/D/1 queue. We then provide upper and lower bounds on the minimum initial buffering required so that the playback interruption probability is below a desired level. The optimal trade-off between the initial buffering and the interruption probability depends on the file size as well as the playback rate compared to the arrival rate of the packets. Our bounds are asymptotically tight as the file size tends to infinity. Moreover, when the arrival rate and the play rate match, we show that the minimum initial buffer size grows as the square-root of the file size. However, if the arrival rate is slightly larger than the play rate, the minimum initial buffering for a given interruption probability remains bounded as the file size grows. In particular, for the infinite file size case, we establish the following relation

$$\text{Probability of interruption} = e^{-I(R)\cdot(\text{initial buffering})}, \qquad (2)$$

where we define $I(R)$ as the *interruption exponent* or reliabil-

† This material is based upon work under subcontract # 18870740-37362-C issued by Stanford University and supported by DARPA. This work is also supported in part by NSF grants CNS 0904520 and CNS 0963818. Initial results were presented in [1].

Ali ParandehGheibi (parandeh@mit.edu), Muriel Médard (medard@mit.edu) and Asuman Ozdaglar (asuman@mit.edu) are with Department of Electrical Engineering and Computer Science, MIT.

S. Shakkottai (sshakkot@tamu.edu) is with the Dept. of Electrical and Computer Engineering,Texas A&M University.

ity function in analogy with (1). In (2), the reliability metric (interruption probability) is related to the delay metric (initial buffering) via the reliability function $I(R)$. We further study the problem of streaming an infinite file size over channels with memory, which are modeled using Markov modulated arrival processes. We establish a relation similar to (2), and characterize the interruption exponent.

There is significant work in the space of P2P streaming. Close to our work, [7], [8], [9], [10] develop analytical models on the trade-off between the steady state probability of missing a block, and buffer size under different block selection policies for live streaming in a full mesh P2P network with deterministic channels. A further modification is to use random linear network coding techniques [11] to make block selection simpler [12], [13], [14], [15] in the wired and wireless context. In contrast, we focus on a very different scenario of streaming of pre-prepared content over unreliable wireless channels using network coding. Further, our analysis is on transient effects—we are interested in the first time that video playback is interrupted as a function of the initial amount of buffering. In another related work Liang and Liang [16] and Luan *et al.* [17] address a similar problem mainly through extensive simulations or diffusion approximation, but their analytical bounds does not provide any further insight on system design in terms of system parameters such as the file size.

For infinite file sizes, the techniques we use to compute the optimal trade-off curves, are similar to those used in the literature of Ruin Theory [18], which study insurer's vulnerability to insolvency. In particular, Reinhard [19] employs a system of integro-differential equations to characterize the non-ruin probabilities of an insurer with constant premium rates and exponentially distributed claim amounts in a Markovian environment. For the finite file size case, we need to characterize hitting probabilities of crossing a time-varying threshold for which such methods are not effective. Our work could be of independent interest since it provides novel techniques for characterizing the trade-offs with finite file sizes.

The rest of this paper is organized as follows. In Section II, we give an overview of a media streaming system with network coding. In Section III, we present the system model and formally define QoE metrics. Section IV is dedicated to characterization of the optimal trade-off curves for Poisson arrivals. In Section V, we generalize these results to case of Markov modulated arrival processes. We verify our analytical results numerically in Section VI. Finally, concluding remarks and extensions are discussed in Section VII.

II. System Overview

We consider a media streaming system as follows. Media files are divided into blocks consisting of multiple frames. The video coding is such that all the frames in the block need to be available before any frames can be played. Blocks are requested in sequence by the playback application from the user-end. The server (or other peers) packetize the requested block and transmit them to the user as in Figure 1. Requesting packets form a single peer may cause delays due to channel uncertainty. However, requesting each packet from multiple

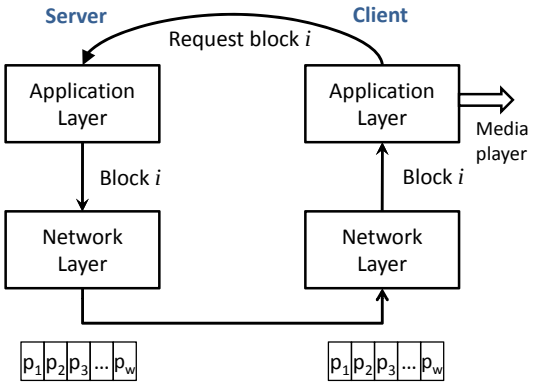

Fig. 1. The media player (application layer) requires complete blocks. At the network layer each block is divided into packets and delivered.

peers introduces the need to keep track of packets, and the duplicate packet reception problem. This is alleviated by using random linear network coding. Here, instead of requesting a particular packet in block i, the receiver simply requests a random linear combination of all the packets in block i. The coefficients of each combination are chosen uniformly at random from a Galois field of size q. The coded packets delivered to the receiver can be thought of as linear equations, where the unknowns are the original packets in block i. Block i can be fully recovered by solving a system of linear equations if it is full rank. It can be shown that if the field size q is large enough, the received linear equations are linearly independent with very high probability [11]. Therefore, for recovering a block of W packets, it is sufficient to receive W coded packets from different peers. In a P2P system, it is unlikely that a randomly contacted peer would have all packets corresponding to a particular block. However, storing blocks in a random linear coded fashion ensures that the selected peer has a useful equation to offer (see [14] for further discussion).

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

III. System Model and QoE Metrics

Consider a single user receiving a media file from various peers it is connected to. Each peer could be a wireless access point or another wireless user operating as a server. We assume that the media file consists of T packets that are divided into blocks of W packets. Each server sends random linear combinations of the packets within the current block to the receiver. We assume that the linear combination coefficients are selected from a Galois field of large enough size. Moreover, we assume that the block size W is small compared to the total length of the file, but large enough to ignore the boundary effects of moving from one block to the next. Time is continuous, and the arrival process of packets from each peer is a Poisson process independent of other arrival processes. Since no redundant packet is delivered

from different peers, we can combine the arrival processes into one Poisson process of rate R. We normalize the playback rate to one, i.e., it takes one unit of time to play a single packet. Thus, our simplified model is just a single-server-single-receiver system. We also assume that the parameter R is known at the receiver, which first buffers D packets from the beginning of the file, and then starts the playback. The dynamics of the receiver's buffer size $Q(t)$ can be described as follows

$$Q(t) = D + A(t) - t, \tag{3}$$

where D is the initial buffer size and $A(t)$ is a Poisson process of rate R. The presence of some packets in the buffer does not guarantee that there will be no interruption since we require W packets corresponding to a block before it can be decoded and played out. However, if there are at least W packets in the buffer, there is at least one playable packet. We declare an interruption in playback when the buffer size decreases to the threshold W. For simplicity of notation, we assume that an *extra* block is initially buffered (not taken into account in D). Hence, we can declare an interruption in playback when the buffer size reaches zero before reaching the end of the file. More precisely, let

$$\tau_e = \inf\{t : Q(t) \leq 0\}, \qquad \tau_f = \inf\{t : Q(t) \geq T - t\}, \tag{4}$$

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

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

$$p(D) = \mathbf{Pr}\{\tau_e < \tau_f\}, \tag{5}$$

where τ_e and τ_f are defined in (4). In our model, the user expects to have an interruption-free experience with probability higher than a desired level $1-\epsilon$. Note that there is a fundamental trade-off between the interruption probability ϵ and the initial buffer size D. For example, owing to the randomness of the arrival process, in order to have zero probability of interruption, it is necessary to fully download the file, i.e., $D = T$. Nevertheless, we need to buffer only a small fraction of the file if user tolerates a positive probability of interruption. These trade-offs and their relation to system parameters R and T are addressed in the following section.

IV. Optimal QoE Trade-offs

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

$$D^*(\epsilon) \quad = \quad \min\{D \geq 0 : p(D) \leq \epsilon\}, \tag{6}$$

where $p(D)$ is the interruption probability defined in (5). Note that in general $p(D)$ and hence $D^*(\epsilon)$ depend on the arrival rate R and the file size T which are assumed to to be known. In the following we characterize the optimal trade-off between the initial buffer size and the interruption probability by providing bounds on $D^*(\epsilon)$. An upper bound (achievability theorem) on $D^*(\epsilon)$ is particularly useful, since it provides a sufficient condition for desirable user experience. A lower bound (converse theorem) of $D^*(\epsilon)$ is helpful to show that the provided upper bound is close to the exact value. The proofs of the main theorems are included in the appendix.

Theorem 1. [Achievability] *Let $D^*(\epsilon)$ be defined as in (6), and $I(R)$ be the largest root of $\gamma(r) = r + R(e^{-r} - 1)$, i.e.,*

$$I(R) = \sup\{r : \gamma(r) = 0\}. \tag{7}$$

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

$$D^*(\epsilon) \leq \frac{1}{I(R)} \log\left(\frac{1}{\epsilon}\right). \tag{8}$$

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

$$\begin{aligned} D^*(\epsilon) \quad \leq \quad & \min\Big\{T(1-R) + \Big(2TR\log\left(\frac{1}{\epsilon}\right)\Big)^{\frac{1}{2}}, \\ & \frac{1}{I(R)}\log\left(\frac{1}{\epsilon}\right)\Big\}. \end{aligned} \tag{9}$$

When the arrival rate R is smaller than one (the playback rate), the upper bound in Theorem 1 consists of two components. The first term, $T(1-R)$, compensates the expected number of packets that are required by the end of $[0, T]$ period. The second component, $\left(2TR\log\left(\frac{1}{\epsilon}\right)\right)^{\frac{1}{2}}$, compensates the randomness of the arrivals to avoid interruptions with high probability. Note that this term increases by decreasing the maximum allowed interruption probability, and it would be zero for a deterministic arrival process. For the case when the arrival rate is larger than the playback rate, the minimum required buffer size does not grow with the file size. This is so since the buffer size in (3) has a positive drift. Hence, if there is no interruption at the beginning of the playback period, it becomes more unlikely to happen later.

In the following, we show that the upper bounds presented in Theorem 1 are *asymptotically tight*, by providing lower bounds on the minimum required buffer size $D^*(\epsilon)$, for different regimes of the arrival rate R. Let us first define the notion of a tight bound.

Definition 1. Let $\hat{D}$ be a lower or upper bound of the $D^*(\epsilon)$, which depends on the file size T. The bound $\hat{D}$ is an *asymptotically tight* bound if $\frac{|\hat{D} - D^*(\epsilon)|}{D^*(\epsilon)}$ vanishes as T goes to infinity.

Theorem 2. [Converse] *Let $D^*(\epsilon)$ be defined as in (6), and $I(R)$ be given by (7). Then*

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

$$D^*(\epsilon) \geq -\frac{1}{I(R)}\log\left(\epsilon + 2e^{-\frac{(R-1)^2}{4(R+1)}T}\right). \tag{10}$$

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

$$D^*(\epsilon) \quad \geq \quad T(1-R) + \frac{1}{2}\Big(2TR\log\left(\frac{1}{\epsilon}\right)\Big)^{\frac{1}{2}} \tag{11}$$

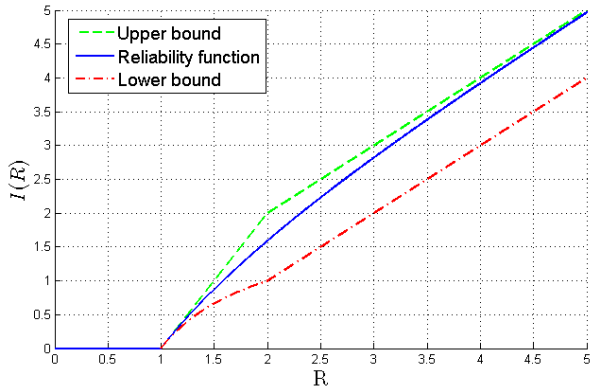

Fig. 2. The reliability function (interruption exponent) defined in (7) for the Poisson arrival process. Simple lower and upper bounds are given by Lemma 1.

Note that the inequality (11) in part (b) of Theorem 2 does not hold for all ϵ. In fact, we can show that $D^*(\epsilon) < T(1-R)$ for a large interruption probability ϵ. In the extreme case $\epsilon = 1$, it is clear that $D^*(\epsilon) = 0$. Nevertheless, since we are interested in *avoiding* interruptions, we do not study this regime of the interruption probabilities. Comparing the lower bounds obtained in Theorem 2 with the upper bounds obtained in Theorem 1, we observe that they demonstrate a similar behavior as a function of the parameters T and R.

Corollary 1. *The upper bounds and lower bounds of $D^*(\epsilon)$ given by Theorems 1 and 2 are asymptotically tight, if $R > 1$, or $R < 1$ and $\epsilon \leq \frac{1}{16}$.*

Corollary 2. *Let $p(D)$ be the interruption probability defined in (5). For the case of $R > 1$, and infinite file size, $T = \infty$, we have*

$$p(D) = \mathbf{Pr}\Big\{ \min_{t\geq 0} Q(t) \leq 0 \Big\} = e^{-I(R)D}, \qquad \text{for all } D \geq 0, \tag{12}$$

where $I(R)$ is defined in (7).

Proof: The proof simply follows from Theorems 1 and 2, and continuity of the probability measure. ∎

Corollary 2 relates the *reliability* of media playback (interruption probability) to *delay* in media playback (initial buffer size) via the *reliability function*, $I(R)$. The reliability function or *interruption exponent* depends on the *rate* and distribution of the arrival process, i.e., properties of the communication channel. Figure 2 plots the reliability function $I(R)$ as a function of the arrival rate. Observe that as the ratio of the arrival rate and the playback rate increases, so does the reliability function. When arrival and playback rates match, the reliability function is zero. This behavior is reminiscent of the error exponent of a noisy communication channel, i.e., the error exponent is zero when the rate is equal to the capacity, and is increasing in the distance between the communication rate and the capacity. In the context of media streaming, Corollary 2 provides an analogue of the error-exponent characterizations that capture the delay-rate-reliability trade-off of block channel codes used for communication over noisy channels.

Thus far, we have studied the QoE trade-offs for media streaming applications with Poisson arrivals. This model does not capture the burstiness of the traffic often associated with correlated losses over a wireless channel. In the following, we generalize the results of this section to the case where packets arrive according to a Markovian process. These results are analogous to error exponent characterization of channels with memory such as the Gilbert-Elliot channel [20].

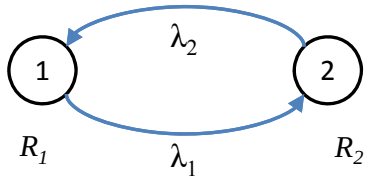

Fig. 3. Two-state Markov process used to model the burstiness of packet arrivals. λ_1 and λ_2 denote transition rates.

V. QoE Trade-offs for Bursty Traffic (Markovian Channels)

Consider a media streaming scenario from a server to a single receiver. In this part, we focus on the case where the media file size is infinite, and the packet arrival rate is strictly larger than the playback rate (normalized to one). The packet arrival process is governed by a two-state continuous time Markov process depicted in Figure 3. For each state i, λ_i denotes the transition rate from state i to the other state. We may verify that the stationary distribution of this Markov process is given by $\pi_i = 1 - \frac{\lambda_i}{\lambda_1+\lambda_2}$, $i = 1, 2$.

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

$$\bar{R} = \sum_{i=1,2} \pi_i R_i = \frac{\lambda_2 R_1 + \lambda_1 R_2}{\lambda_1 + \lambda_2} > 1. \tag{13}$$

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

A. Markovian Channels with Deterministic Arrivals

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

Theorem 3. *Consider the Markov process depicted in Figure 3. Let the arrival process at the receiver be deterministic with rate R_i, when in state i. If $R_1 > 1 > R_2$ and (13) holds, then*

$$p_1(D) = \frac{\rho_1}{\rho_2} e^{-(\rho_2-\rho_1)D}, \quad p_2(D) = e^{-(\rho_2-\rho_1)D}, \tag{14}$$

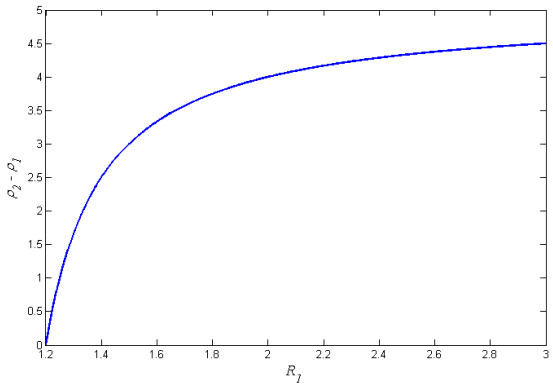

Fig. 4. The reliability function, $\rho_2 - \rho_1$, given by Theorem 3 for the Gilbert-Elliot channel with deterministic arrivals. Here, $\lambda_1 = \lambda_2 = 1$, and $R_2 = 0.8$.

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

Proof: Since we are considering the infinite file size case, the interruption probability starting from any particular time, only depends on the queue-length and the state of the Markov process at that time. For any $D, h \geq 0$, write

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

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

$$\frac{\partial p_i(D)}{\partial D} = \rho_i\big(p_1(D) - p_2(D)\big), \quad i = 1, 2, \tag{15}$$

with the boundary condition $p_i(\infty) = 0$. The assumption in (13) is necessary for this boundary condition to hold. Using (15) together with the boundary condition, we have $p_1(D) = \frac{\rho_1}{\rho_2} p_2(D)$. Replacing this relation back in (15) for $i = 2$, we obtain a differential equation for $p_2(D)$

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

which has the solution $p_2(D) = c_1 e^{-(\rho_2 - \rho_1)D} + c_2$. Observe that $c_2 = 0$ by the boundary condition. Since $R_2 < 1$, we also have $p_2(0) = 1$. Hence, $c_1 = 1$, that proves the claim in (14). ∎

We observe similar exponential decay in the interruption probability as in Corollary 2. The interruption exponent (reliability function) given by $\rho_2 - \rho_1$ is increasing in R_1 and R_2. Figure 4 plots the reliability function, $\rho_2 - \rho_1$, as a function of R_1, when $R_2 = 0.8$ and $\lambda_1 = \lambda_2 = 1$. Observe that the reliability function goes to zero as the average arrival rate $\bar{R}$ approaches the playback rate. It is worth mentioning that for a fixed average arrival rate, as λ_1 and λ_2 increase the mixing time of the Markov process decreases and the hence, the arrival process tend to look deterministic with rate $\bar{R} > 1$. Therefore, larger reliability function is achieved. This behavior is similar to that of the conventional Gilbert-Elliot communication channel [20].

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

B. Markovian Channels with Poisson Arrivals

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

Theorem 4. *Consider the two-state Markov process depicted in Figure 3. Let the arrival process at the receiver be Poisson with rate R_i, when in state i. If the average arrival rate is larger than the playback rate, i.e., (13) holds, then*

$$p_1(D) = c_{11} e^{-s_1 D} + c_{12} e^{-s_2 D}, \; p_2(D) = c_{21} e^{-s_1 D} + c_{22} e^{-s_2 D}, \tag{16}$$

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

$$\Phi(s) = (s + R_1(e^{-s} - 1) - \lambda_1)(s + R_2(e^{-s} - 1) - \lambda_2) - \lambda_1 \lambda_2, \tag{17}$$

and

$$\begin{aligned} c_{i1} &= \frac{s_2 + R_i(e^{-s_2} - 1)}{(s_2 - s_1) + R_i(e^{-s_2} - e^{-s_1})}, \\ c_{i2} &= -\frac{s_1 + R_i(e^{-s_1} - 1)}{(s_2 - s_1) + R_i(e^{-s_2} - e^{-s_1})}. \end{aligned} \tag{18}$$

Proof: We sketch the proof owing to space limitation. Similarly to the proof of Theorem 3, and [19], we obtain the following system of delay differential equations

$$\begin{aligned} \frac{\partial p_i(D)}{\partial D} \quad &= \quad R_i\Big(p_i(D+1) - p_i(D)\Big) \\ &\quad + \lambda_i\Big(p_{-i}(D) - p_i(D)\Big), \quad i = 1, 2. \end{aligned} \tag{19}$$

Using the assumption in (13), we get the boundary conditions $p_1(\infty) = p_2(\infty) = 0$. Moreover, since the packets depart from the receiver's queue deterministically, we obtain the additional boundary conditions $p_1(0) = p_2(0) = 1$. We may solve the above system of differential equations using the roots of characteristic equation given by (17). Observe that the first set of boundary conditions imply that we only need to consider the terms that vanish as D grows to infinity. We may verify that $\Phi(s)$ has two roots $s_1, s_2 > 0$ if the condition in (13) holds (see Figure 5). Therefore, the interruption probabilities take the form of (16). Further, the second set of boundary conditions imply that $c_{i1} + c_{i2} = 1$ for $i = 1, 2$. We may solve for the coefficients in (18) by plugging (16) in (19), and using these boundary conditions. ∎

The behavior of the interruption probabilities given by Theorem 4 is governed by the dominant mode s_1. Even though, we cannot compute the roots of the characteristic function analytically, we may examine approximate solutions of the characteristic equation when the Markov process is fast

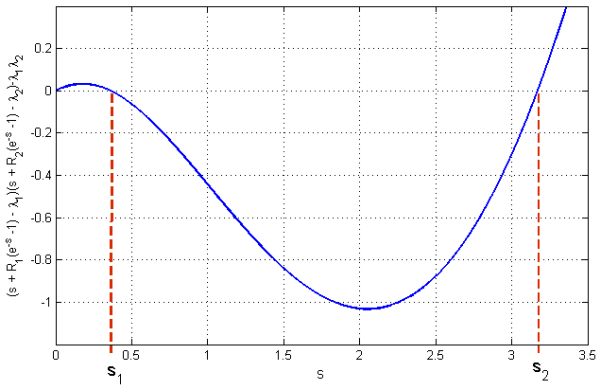

Fig. 5. The characteristic function $\Phi(s)$ given by (17). s_1 denotes the dominant root.

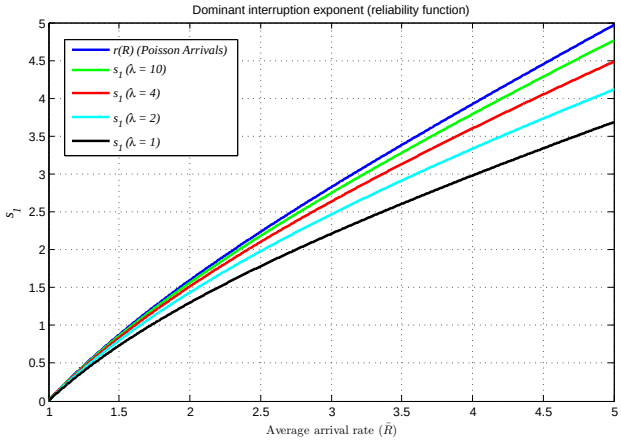

Fig. 6. The dominant root of the characteristic function given by (17), as a function of the average arrival rate for different $\lambda_1 = \lambda_2 = \lambda$. Here, we select $R_1 = 1.4\bar{R}$, and $R_2 = 0.6\bar{R}$.

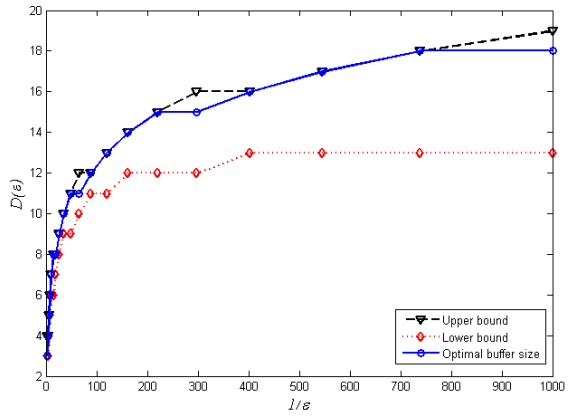

Fig. 7. The minimum buffer size $D^*(\epsilon)$ as a function of the interruption probability. The perfect match between the optimal curve and the upper bound is due to the rounding effect.

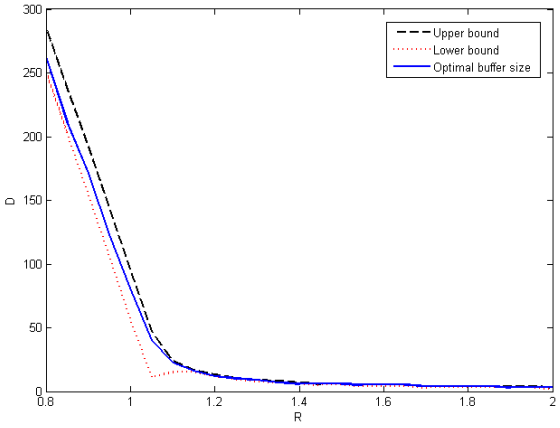

Fig. 8. The minimum buffer size $D^*(\epsilon)$ as a function of the arrival rate R.

mixing, i.e., $\lambda_1, \lambda_2 \gg 1$. Using the notation in (13), we can rewrite (17) as

$$\begin{aligned}\Phi(s) &= \big(s + R_1(e^{-s}-1)\big)\big(s + R_2(e^{-s}-1)\big) \\ &\quad -(\lambda_1+\lambda_2)\big(s + \bar{R}(e^{-s}-1)\big). \qquad (20)\end{aligned}$$

Now, let λ_1, λ_2 tend to infinity, while keeping R_1 and R_2 fixed. There are two possibilities for the roots of $\Phi(s)$: First, the root does not grow to infinity; second, the root scales to infinity as λ_1, λ_2 grow. In the former case, we need to have the coefficient of $(\lambda_1 + \lambda_2)$ in (20) go to zero, i.e., for λ_1, λ_2 large enough we have $s + \bar{R}(e^{-s} - 1) \approx 0$. Therefore, using the notation in (7), $s = I(\bar{R})$ is an approximate root of $\Phi(s)$. Figure 6 plots s_1, the exact root of $\Phi(s)$, as a function of the average arrival rate, $\bar{R}$, for different mixing times. Observe that as λ grows, s_1 approaches $I(\bar{R})$, which is the reliability function for the Poisson arrival case (see Corollary 2). In the second case, we may approximate the characteristic function in (20) as $\Phi(s) \approx s^2 - (\lambda_1+\lambda_2)s$. This results in $s_2 \approx \lambda_1 + \lambda_2$ that grows to infinity with λ_1 and λ_2. This root is significantly larger than the dominant root, s_1, which remains bounded. Finally, note that if $R_1 = R_2 = R$, we have $c_{12} = c_{22} = 0$ for all λ_1 and λ_2. This confirms the earlier result of Corollary 2.

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

VI. Numerical Results

We use MATLAB simulations to compute the minimum initial buffer size $D^*(\epsilon)$ for a given interruption probability ϵ in various scenarios. We start from a small initial buffer size D, and for each D we compute the interruption probability $p(D)$ via the Monte-Carlo method. We increase D until the constraint $p(D) \leq \epsilon$ is satisfied. Since $p(D)$ is monotonically decreasing in D, this gives the minimum required buffer size. We restrict D to take only integer values, and round each upper bound value up, and each lower bound value down to the nearest integer.

Figure 7 shows the minimum required buffer size $D^*(\epsilon)$ as a function of $\frac{1}{\epsilon}$, as well as the bounds given by Theorems 1, 2. The arrival rate is fixed to $R = 1.2$ and the file size $T = 500$. We observe that the numerically computed trade-off curve closely matches our analytical results.

Figure 8 plots the minimum required buffer size $D^*(\epsilon)$ as well as the upper and lower bounds given by Theorems 1 and 2 versus the arrival rate R, where $\epsilon = 10^{-2}$ and the file size is fixed to $T = 10^3$. Note that when the arrival rate is almost equal or less than the playback rate, increasing the arrival rate can significantly reduce the initial buffering delay. However, for larger arrival rates $D^*(\epsilon)$ is small enough such that increasing R does not help anymore.

VII. Conclusions

In this paper, we studied the problem of media streaming with focus on the trade-offs between the two QoE metrics—probability of interruption in media playback, and the initial waiting time before starting the playback. In our system, the user can receive packets of the media stream from multiple sources by requesting packets in each block of the file. We used the fact that sending random linear combinations of the packets within each block of the media file simplifies the packet selection strategies of P2P systems. This fact allowed us to describe the receiver's buffer dynamics as an M/D/1 queue, and explicitly characterize the trade-off between the QoE metrics for different ranges of the system parameters.

We observed that the minimum initial buffer size to attain a desired level of interruption probability scales as the square root of the file size if the arrival rate and the playback rate match. However, when the arrival rate is slightly larger than the playback rate, the initial buffer size remains bounded. Moreover, we proved that our bounds on the optimal trade-off curve are asymptotically tight as the file size grows. We further generalized the results to the channels with memory modeled using Markov modulated arrival processes. We explicitly characterized interruption probabilities for two-state Markov processes. The generalization to any finite state Markov modulated process is straightforward. Finally, our numerical results confirmed that the optimal trade-off curves demonstrate a similar behavior to that predicted by our bounds.

This work is the first step in analytical characterization of QoE trade-offs in wireless media streaming applications. Many of the insights obtained from the presented simple model may be extended to a general class of arrival processes. Other future research directions include further justification of our system model through extensive simulations using real network traces, generalization of the QoE metrics to allow for few *tolerable* interruptions during media playback, and characterization of the optimal *re-buffering* policies once interruptions occur. Another interesting extension to this work would be to obtain optimal resource allocation policies to satisfy users who have different interruption probability and initial waiting time targets.

Appendix

We first prove some useful lemmas used in the proof of Theorem 1.

Lemma 1. *Let $I(R)$ be given by (7). The following relations hold:*

$$I(R) = 0, \qquad \text{if } 0 \le R \le 1, \tag{21}$$

$$\frac{2(R-1)}{R} \le I(R) \le 2(R-1), \qquad \text{if } 1 \le R \le 2, \tag{22}$$

$$R - 1 \le I(R) \le R \le 2(R-1), \qquad \text{if } R \ge 2. \tag{23}$$

Proof: We consider three different cases separately.

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

Case II ($1 \le R \le 2$): By definition of $I(R)$ in (7),

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

Rearranging the terms in the above relation, gives the lower bound in (22). We show the upper bound in two steps. First, we show that $\gamma(2(R-1)) > 0$ for $R > 1$, then we verify that $\gamma(r) \ge 0$ for all $r \ge 2(R-1)$. These two facts imply that $\gamma(r) > 0$ for all $r \ge 2(R-1)$, i.e., $I(R) \le 2(R-1)$. The first step can be verified by noting that

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

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

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

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

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

Lemma 2. *Let $X(t) = e^{-rQ(t)}$, where $Q(t)$ is given by (3). Also let $\gamma(r) = r + R(e^{-r} - 1)$. Then for every $r \ge 0$ such that $\gamma(r) \ge 0$, $X(t)$ is a sub-martingale with respect to the canonical filtration $\mathcal{F}_t = \sigma(X(s), 0 \le s \le t)$, i.e., the smallest σ-field containing the history of the stochastic process X up to time t. Moreover, if $\gamma(r) = 0$ then $X(t)$ is a martingale.*

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

$$\mathbf{E}[X(t+h)|\mathcal{F}_t] \ge X(t) \quad \text{a.s.} \tag{24}$$

$X(t)$ is a martingale if (24) holds with equality. The left-hand side of (24) can be expressed as

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

where (a) follows from independent increment property of the Poisson process, and (b) follows from the fact that $A(t)$ is a Poisson random variable. Now, it is immediate to verify (24) for any r with $\gamma(r) \geq 0$. Finally, note that if $\gamma(r) = 0$, the equality in the above relations hold through, and (24) holds with equality. Therefore, $X(t)$ is a martingale for r with $\gamma(r) = 0$. ∎

Next, we use Doob's maximal inequality [21] to bound the interruption probability.

Lemma 3. *Let $p(D)$ be the interruption probability defined in (5), and $\gamma(r) = r + R(e^{-r} - 1)$. Then, for any $r \geq 0$ with $\gamma(r) \geq 0$*

$$p(D) \leq e^{-rD+T\gamma(r)}, \quad \textit{for all } D, T, R \geq 0. \tag{25}$$

Proof: By definition of $p(D)$ in (5), we have

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

where (a) holds by applying Doob's maximal inequality [21] to the non-negative sub-martingale $X(t) = e^{-rQ(t)}$. Note that $X(t)$ is a sub-martingale for all r with $\gamma(r) \geq 0$ by Lemma 2. ∎

Lemma 4. *It holds that $-(1-z)\log(1-z) - z \leq -\frac{z^2}{2}$, for all $0 \leq z < 1$.*

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

Proof of Theorem 1. First, note that for any upper bound $\bar{p}(D)$ of the interruption probability $p(D)$, any feasible solution of

$$\bar{D}(\epsilon) = \min\{D \geq 0 : \bar{p}(D) \leq \epsilon\} \tag{26}$$

provides an upper bound on $D^*(\epsilon)$. This is so since the optimal solution of the above problem is feasible in the minimization problem (6). If the problem in (26) is infeasible, we use the convention $\bar{D}(\epsilon) = \infty$, which is a trivial bound on $D^*(\epsilon)$. The rest of the proof involves finding the tightest bounds on $p(D)$ and solving (26).

Part (a): By Lemma 3, for $r = I(R)$, we can write $p(D) \leq \bar{p}_a(D) = e^{-I(R)D}$, for all $D, T, R \geq 0$. Solving $\bar{p}_a(D) = \epsilon$ for D gives the result of part (a). Since $I(R) = 0$ for $R \leq 1$ (cf. Lemma 1), this bound is not useful in that range.

Part (b): First, we claim that for all $D \geq T(1-R+I(R))$, we have $p(D) \leq \bar{p}_b(D) = e^{-\frac{1}{2}TRz^2}$, where $z = 1 - \frac{1}{R}\big(1 - \frac{D}{T}\big)$. We use Lemma 3 with $r = r^* = -\log\big(\frac{1}{R}\big(1 - \frac{D}{T}\big)\big)$ to prove the claim. Note that $r^* \geq 0$, because $D \geq T(1-R)$. In order to verify the second hypothesis of Lemma 3, consider the following

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

where (a) and (b) follow from the definition of $I(R)$ and r^*, respectively, and (c) holds by the hypothesis of the claim. Thus, $r^* \geq I(R)$. Using the facts that $I(R)$ is the largest root of $\gamma(r)$, and $\gamma(r) \to +\infty$ as $r \to \infty$, we conclude that $\gamma(r^*) \geq 0$. Now, we apply Lemma 3 to get

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

where (a) and (b) follow from the definition of $\gamma(r)$ and z, and (c) is true by Lemma 4. Therefore, the claim holds.

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

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

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

Next, we establish some lemmas used in the proof of Theorem 2. The proofs of Lemmas 5 and 6 are based on the results from [22], [23], and are removed owing to space limitations.

Lemma 5. *Let Z be a Poisson random variable with mean λ. If $\lambda \geq 2$, and $k \geq 2$, then*

$$\mathbf{Pr}\{Z \leq \lambda - k\} \leq \exp\Big(\frac{1}{2\lambda}\big(k - \frac{3}{2}\big)^2\Big). \tag{27}$$

Lemma 6. *Let Z be a Poisson random variable with mean λ. For all $m \leq \frac{\sqrt{\lambda}}{20} - 1$*

$$\mathbf{Pr}\{Z \leq \lambda - m\sqrt{\lambda}\} \geq \frac{1}{3}\exp\Big(-\frac{1}{1.9}\big(m + \frac{1}{2}\big)^2\Big). \tag{28}$$

Lemma 7. *Let $Q(t)$ be given by (3). For $s, \delta > 0$, define the boundary function*

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

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

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

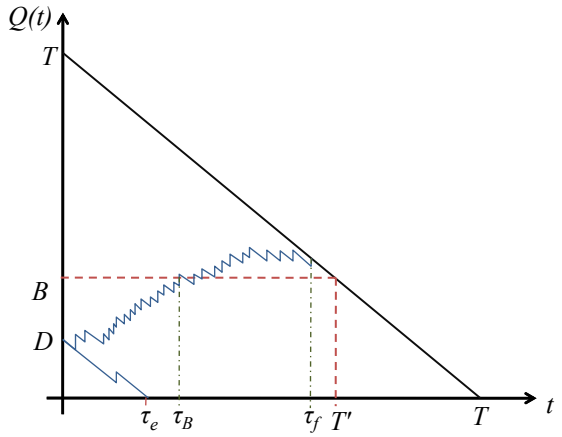

Fig. 9. Two sample paths of the buffer size $Q(t)$ demonstrating the interruption event at time τ_e, crossing the threshold B at time τ_B, and the download complete event at time τ_f.

Proof: Define $Z(t) = e^{rQ'(t)}$, where $Q'(t) = Q(t) - u(t) = -\delta + A(t) + (R+s)t$, and $r > 0$ satisfies $\varphi(r) := e^r - 1 - r(1+\frac{s}{R}) = 0$. Similarly to the proof of Lemma 2, we can show that $Z(t)$ is a martingale. This allows us to use Doob's maximal inequality to obtain

$$\begin{aligned}\mathbf{Pr}\Big\{\sup_{0\le t\le T} Q(t) \ge u(t)\Big\} &= \mathbf{Pr}\Big\{\sup_{0\le t\le T} Z(t) \ge 1\Big\}\\ &\le \mathbf{E}[Z(T)] = e^{-r\cdot\delta}.\end{aligned}$$

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

$$\begin{aligned}\varphi(\frac{s}{R}) &= e^{\frac{s}{R}} - 1 - \frac{s}{R}(1+\frac{s}{R})\\ &\le \frac{s}{R} + (\frac{s}{R})^2 - \frac{s}{R}(1+\frac{s}{R}) = 0. \qquad (29)\end{aligned}$$

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

Proof of Theorem 2. *Part(a):* Similarly to the argument as in the proof of Theorem 1, it is sufficient to provide a lower bound on $p(D)$ defined in (5).

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

A necessary condition for the interruption event to happen is to have the receiver's buffer emptied before time $T' = T - B$, or crossing the threshold B (see Figure 9). In particular,

$$p(D) = \mathbf{Pr}\{\tau_e < \tau_f\} \ge \mathbf{Pr}\big\{\tau_e \le \min\{\tau_B, T'\}\big\}. \qquad (30)$$

Define the stopping time $\tau = \min\{\tau_e, \tau_B, T'\}$, and let $Y(t) = e^{-I(R)Q(t)}$, where $I(R) > 0$ is given by (7). By Lemma 2, $Y(t)$ is a martingale. Moreover, $Y(t) \le 1$, and $\tau \le T < \infty$. Therefore, we can apply Doob's optional stopping theorem [21] to get

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

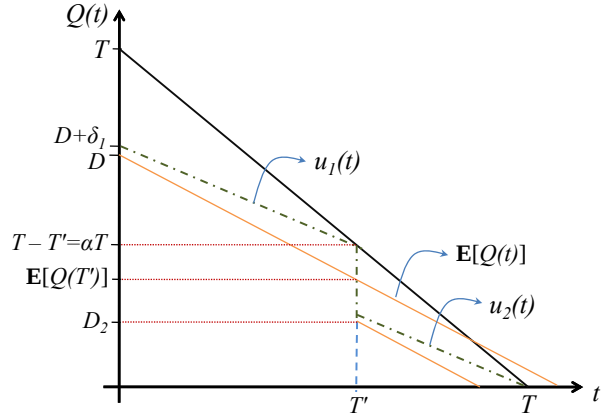

Fig. 10. Guideline for the proof of Theorem 2(b).

where (a) is the result of Doob's optional stopping time. (b) holds because $Y(t) \le 1$ for all t, and $Y(t) \le e^{-I(R)\cdot B}$ if $Q(t) \ge B$. Finally, (c) follows from (30). Rearranging the terms in the above relation, we obtain

$$p(D) \ge e^{-I(R)D} - e^{-I(R)B} - \mathbf{Pr}\{\tau = T'\}. \qquad (31)$$

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

$$\begin{aligned}\mathbf{Pr}\{\tau = T'\} &\overset{(a)}{\le} \mathbf{Pr}\{0 \le Q(T-B) \le B\}\\ &\overset{(b)}{=} \mathbf{Pr}\{\alpha T - D \le A(\alpha T) \le T - D\}\\ &\le \mathbf{Pr}\big\{A(\alpha T) \le R\alpha T - ((R\alpha-1)T + D)\big\}\\ &\overset{(c)}{\le} \exp\left(-\frac{1}{2R\alpha T}\big((R\alpha-1)T + D - \frac{3}{2}\big)^2\right)\\ &\overset{(d)}{\le} \exp\left(-\frac{(R-1)^2}{4(R+1)}T\right), \qquad (32)\end{aligned}$$

where (a) holds because $Q(t)$ cannot be negative or above the threshold B if stopping at T', and (b) follows from the buffer dynamics in (3). Recall that $A(\alpha T)$ is a Poisson random variable with mean $R\alpha T$. Since $\alpha \ge \frac{1}{R}$, (c) holds for $D, T \ge 2$ by employing Lemma 5 with

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

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

By Lemma 1, we have $I(R) \ge (R-1)$ for all $R > 1$. Therefore, we can bound the second term in (31) as follows

$$\begin{aligned}e^{-I(R)B} &\le \exp\left(-(R-1)\frac{R-1}{2R}T\right)\\ &\le \exp\left(-\frac{(R-1)^2}{4(R+1)}T\right). \qquad (33)\end{aligned}$$

Combine the bounds in (32) and (33) with (31) to obtain

$$p(D) \ge e^{-I(R)D} - 2e^{-\frac{(R-1)^2}{4(R+1)}T}, \quad \text{for all } D, T \ge 2. \qquad (34)$$

Therefore, $p(D) \ge \epsilon$ if $D = -\frac{1}{I(R)}\log\left(\epsilon + 2e^{-\frac{(R-1)^2}{4(R+1)}T}\right) \ge 2$. This immediately gives the result in (10). For the case in which $D < 2$ or $T < 2$, the claim holds trivially.

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

$$D = T(1-R) + \frac{1}{2}\Big(2TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}. \qquad (35)$$

Let us first define the boundary functions

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

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

$$\delta_1 = \frac{1}{2}\alpha TR-\frac{1}{4}\Big(2TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}, \tag{36}$$
$$\delta_2 = \Big(\alpha TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}, \tag{37}$$
$$D_2 = \alpha T(1-R)-2\Big(\alpha TR\log\big(\frac{1}{\epsilon}\big)\Big)^{\frac{1}{2}}. \tag{38}$$

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

Observe that for every sample path of the buffer size $Q(t)$, the only way for completing the file download, $Q(t)=T-t$, is to cross the boundary functions $u_i(t)$ or have $Q(T')\geq D_2$ (cf. Figure 10). This gives a necessary condition for the interruption event. Hence,

$$\begin{aligned}
p(D) &= \mathbf{Pr}\{\tau_e<\tau_f\}\geq\mathbf{Pr}\{\tau_e\leq\min\{\tau_1,T'\}\}\\
&\quad+\mathbf{Pr}\{\tau_e\leq\tau_2,T'\leq\min\{\tau_e,\tau_1\},Q(T')\leq D_2\}\\
&= \mathbf{Pr}\{\tau_e\leq\min\{\tau_1,T'\}\}\\
&\quad+\mathbf{Pr}\{\tau_e\leq\tau_2|T'\leq\min\{\tau_e,\tau_1\},Q(T')\leq D_2\}\big[1\\
&\quad-\mathbf{Pr}\{\tau_1\leq\min\{\tau_e,T'\}\}-\mathbf{Pr}\{\tau_e\leq\min\{\tau_1,T'\}\}\\
&\quad-\mathbf{Pr}\{T'\leq\min\{\tau_e,\tau_1\},Q(T')>D_2\}\big]\\
&\geq \mathbf{Pr}\{\tau_e\leq\tau_2|T'\leq\min\{\tau_e,\tau_1\},Q(T')\leq D_2\}\big[1\\
&\quad-\mathbf{Pr}\{\tau_1\leq\min\{\tau_e,T'\}\}\\
&\quad-\mathbf{Pr}\{T'\leq\min\{\tau_e,\tau_1\},Q(T')>D_2\}\big].
\end{aligned}\tag{39}$$

In the following we provide bounds on each of the terms in (39). By Markov property of the Poisson process we have

$$\begin{aligned}
&\mathbf{Pr}\{\tau_e\leq\tau_2|T'\leq\min\{\tau_e,\tau_1\},Q(T')\leq D_2\}\\
&=\mathbf{Pr}\{\tau_e\leq\tau_2|Q(T')\leq D_2\}\\
&\geq\mathbf{Pr}\{\tau_e\leq\tau_2|Q(T')=D_2\}\\
&=1-\mathbf{Pr}\{\tau_2\leq\tau_e|Q(T')=D_2\}\\
&\geq 1-\mathbf{Pr}\Big\{\sup_{T'\leq t\leq T}Q(t)\geq u_2(t)|Q(T')=D_2\Big\}\\
&\geq 1-e^{-\frac{\delta_2^2}{R(T-T')}}=1-\epsilon,
\end{aligned}\tag{40}$$

where the last inequality follows from Lemma 7 with parameters $\delta=\delta_2$ and $s=\frac{\delta_2}{T-T'}$, if its hypothesis $s\leq R$ is satisfied. This is equivalent to having T satisfy $T\geq\frac{1}{\alpha R}\log\big(\frac{1}{\epsilon}\big)$.

Similarly, by employing Lemma 7 with $\delta=\delta_1$ and $s=\frac{\delta_1}{T'}$ we have

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

Note that the hypothesis of Lemma 7 is satisfied here for all $\alpha\leq\frac{2}{3}$. Moreover, if $T\geq\frac{16}{\alpha^2R}\log\big(\frac{1}{\epsilon}\big)$, we have

$$\begin{aligned}
\mathbf{Pr}\{\tau_1\leq\min\{\tau_e,T'\}\} &\leq \exp\Big(-4\log\big(\frac{1}{\epsilon}\big)[1-\frac{1}{\sqrt{32}}]^2\Big)\\
&\leq \exp\Big(-2\log\big(\frac{1}{\epsilon}\big)\Big)=\epsilon^2.
\end{aligned}\tag{41}$$

For the last term in (39) write

$$\begin{aligned}
&1-\mathbf{Pr}\{T'\leq\min\{\tau_e,\tau_1\},Q(T')>D_2\}\\
&\geq 1-\mathbf{Pr}\{Q(T')>D_2\}=\mathbf{Pr}\{Q(T')\leq D_2\}\\
&\quad=\mathbf{Pr}\{D+A(T')-T'\leq D_2\}\\
&=\mathbf{Pr}\{A(T')\leq T'R-m\sqrt{T'R}\},
\end{aligned}\tag{42}$$

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

$$m=\Big(\frac{\log\big(\frac{1}{\epsilon}\big)}{2(1-\alpha)}\Big)^{\frac{1}{2}}\big[1+(8\alpha)^{\frac{1}{2}}\big]. \tag{43}$$

For $T\geq\frac{16}{\alpha^2R}\log\big(\frac{1}{\epsilon}\big)$ and $\alpha,\epsilon\leq\frac{1}{16}$, we may verify that $m\leq\sqrt{RT'}/20-1$. Hence, we can use Lemma 6 to bound (42) from below and conclude

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

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

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

By continuity of m in α (cf. (43)), we can choose $\alpha=\alpha_0>0$ small enough such that $\frac{1}{3}e^{-\frac{1}{1.9}\left(m+\frac{1}{2}\right)^2}\geq\frac{17}{15}\epsilon$. Now, by plugging this relation as well as (40) and (41) back in (39) we have for all $\epsilon\leq\frac{1}{16}$,

$$p(D)\geq(1-\epsilon)\big(\frac{17}{15}\epsilon-\epsilon^2\big)\geq\frac{15}{16}\big(\frac{17}{15}-\frac{1}{16}\big)\epsilon\geq\epsilon, \tag{44}$$

if $T\geq\frac{16}{\alpha_0^2R}\log\big(\frac{1}{\epsilon}\big)$. Therefore, the buffer size D, defined in (35), is a lower bound on $D^*(\epsilon)$. ■

REFERENCES

[1] A. ParandehGheibi, M. Médard, S. Shakkottai, and A. Ozdaglar. Avoiding interruptions - QoE trade-offs in block-coded streaming media applications. in proc. ISIT 2010, arXiv:1001.1937 [cs.MM].

[2] C. Fraleigh, S. Moon, B. Lyle, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and C. Diot. Packet-level traffic measurements from the Sprint IP backbone. *IEEE Network Magazine*, 17(6):6–16, 2003.

[3] K. P. Gummadi, R. J. Dunn, S. Saroiu, S. D. Gribble, H. M. Levy, and J. Zahorjan. Measurement, modeling, and analysis of a peer-to-peer file-sharing workload. In *Proc. SOSP*.

[4] C. Labovitz, D. McPherson, and S. Iekel-Johnson. 2009 Internet Observatory report. In *NANOG-47*, October 2009.

[5] PPLive. http://www.pplive.com/, 2009.

[6] QQLive. http://www.qqlive.com/, 2009.

[7] Y. Zhou, D. Chiu, and J. Lui. A simple model for analyzing P2P streaming protocols. In *Proc. IEEE ICNP 2007*.

[8] T. Bonald, L. Massoulié, F. Mathieu, D. Perino, and A. Twigg. Epidemic live streaming: optimal performance trade-offs. *SIGMETRICS Perform. Eval. Rev.*, 36(1):325–336, 2008.

[9] Bridge Q. Zhao, John C.S. Lui, and Dah-Ming Chiu. Exploring the optimal chunk selection policy for data-driven P2P streaming systems. In *The 9th International Conference on Peer-to-Peer Computing*, 2009.

[10] L. Ying, R. Srikant, and S. Shakkottai. The Asymptotic Behavior of Minimum Buffer Size Requirements in Large P2P Streaming Networks. In *Proc. of the Information Theory and Applications Workshop*, San Diego, CA, 2010.

[11] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger. A random linear network coding approach to multicast. *IEEE Transactions on Information Theory*, 52:4413–4430, 2006.

[12] S. Acedanski, S. Deb, M. Médard, and R. Koetter. How good is random linear coding based distributed networked storage. In *Proc. NetCod*, 2005.

[13] C. Gkantsidis, J. Miller, and P. Rodriguez. Comprehensive view of a live network coding P2P system. In *Proc. ACM SIGCOMM*, 2006.

[14] M. Wang and B. Li. R^2: Random push with random network coding in live peer-to-peer streaming. *IEEE JSAC, Special Issue on Advances in Peer-to-Peer Streaming Systems*, 25:1655–1666, 2007.

[15] H. Chi and Q. Zhang. Deadline-aware network coding for video on demand service over P2P networks. *Journal of Zhejiang University - Science A*, 7(5):755–763, 2006.

[16] G. Liang and B. Liang. Effect of delay and buffering on jitter-free streaming over random VBR channels. *IEEE Transactions on Multimedia*, 10(6):1128–1141, 2008.

[17] T. H. Luan, L. X. Cai, and X. Shen. Impact of network dynamics on users video quality: Analytical framework and QoS provision. *IEEE Transactions on Multimedia*, 12(1):54–78, 2010.

[18] S. Asmuseen. *Ruin Probabilities*. World Scientific Publishing Company, 2000.

[19] J. M. Reinhard. On a class of semi-Markov risk models obtained as classical risk models in a Markovian environment. *ASTIN Bull*, 14(1):23–43, 1984.

[20] R. G. Gallager. *Information Theory and Reliable Communication*. Wiley, 1968.

[21] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1997.

[22] P. W. Glynn. Upper bounds on Poisson tail probabilities. *Operations Research Letters*, 6(1):9–14, 1987.

[23] B. L. Fox and P. W. Glynn. Computing Poisson probabilities. *Communications of the ACM*, 31(4):440–445, 1988.

Ali ParandehGheibi is currently pursuing a Ph.D. degree in Electrical Engineering and Computer Science at the Massachusetts Institute of Technology (MIT), Cambridge, USA. He received a B.Sc. degree (first-class honors) in Electrical Engineering in 2006 from Sharif University of Technology, Tehran, Iran; and a S.M. in Electrical Engineering and Computer Science in 2008 from Massachusetts Institute of Technology, Cambridge. His research interests include network optimization and control, network coding, video streaming, Peer-to-Peer and heterogeneous networks, distributed optimization, gossip algorithms and social networks.

Srinivas Shakkottai (S '00-M '08) received his Bachelor of Engineering degree in electronics and communication engineering from the Bangalore University, India, in 2001 and his M.S. and Ph.D degrees from the University of Illinois at Urbana-Champaign in 2003 and 2007, respectively, both in electrical engineering. He was Postdoctoral Scholar at Stanford University until December 2007, and is currently an Assistant Professor at the Dept. of ECE, Texas A&M University. He has received the Defense Threat reduction Agency Young Investigator Award (2009) as well as research awards from Cisco (2008) and Google (2010).

His research interests include peer-to-peer systems, pricing approaches to network resource allocation, game theory, congestion control, and the measurement and analysis of Internet data.

Muriel Médard is a Professor in the Electrical Engineering and Computer Science Department at the Massachusetts Institute of Technology. She was previously an Assistant Professor in the Electrical and Computer Engineering Department and a member of the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. From 1995 to 1998, she was a Staff Member at MIT Lincoln Laboratory in the Optical Communications and the Advanced Networking Groups. Professor Médard received B.S. degrees in EECS and in Mathematics in 1989, a B.S. degree in Humanities in 1990, a M.S. degree in EE in 1991, and a Sc. D. degree in EE in 1995, all from the Massachusetts Institute of Technology (MIT), Cambridge. She serves as an Associate Editor for the Optical Communications and Networking Series of the IEEE Journal on Selected Areas in Communications, as an Associate Editor in Communications for the IEEE Transactions on Information Theory and as a Guest Editor for the Joint special issue of the IEEE Transactions on Information Theory and the IEEE/ACM Transactions on Networking on Networking and Information Theory. She has served as a Guest Editor for the IEEE Journal of Lightwave Technology and as an Associate Editor for the OSA Journal of Optical Networking.

Professor Médards research interests are in the areas of network coding and reliable communications, particularly for optical and wireless networks. She was awarded the IEEE Leon K. Kirchmayer Prize Paper Award 2002 for her paper, The Effect Upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel, IEEE Transactions on Information Theory, Volume 46 Issue 3, May 2000, Pages: 935946. She was co-awarded the Best Paper Award for G. Weichenberg, V. Chan, M. Médard, Reliable Architectures for Networks Under Stress, Fourth International Workshop on the Design of Reliable Communication Networks (DRCN 2003), October 2003, Banff, Alberta, Canada. She received a NSF Career Award in 2001 and was a co-winner of the 2004 Harold E. Edgerton Faculty Achievement Award, established in 1982 to honor junior faculty members for distinction in research, teaching and service to the MIT community. She was named a 2007 Gilbreth Lecturer by the National Academy of Engineering. Professor Médard is a Fellow of IEEE.

Asu Ozdaglar received the B.S. degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 1996, and the S.M. and the Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1998 and 2003, respectively.

Since 2003, she has been a member of the faculty of the Electrical Engineering and Computer Science Department at the Massachusetts Institute of Technology, where she is currently the Class of 1943 Associate Professor. She is also a member of the Laboratory for Information and Decision Systems and the Operations Research Center. Her research interests include optimization theory, with emphasis on nonlinear programming and convex analysis, game theory, with applications in communication, social, and economic networks, and distributed optimization and control. She is the co-author of the book entitled "Convex Analysis and Optimization" (Athena Scientific, 2003).

Professor Ozdaglar is the recipient of a Microsoft fellowship, the MIT Graduate Student Council Teaching award, the NSF Career award, and the 2008 Donald P. Eckman award of the American Automatic Control Council. She currently serves as an associate editor for the area Optimization Theory, Algorithms and Applications for The Asia-Pacific Journal of Operational Research and is a member of the Board of Governors of the Control System Society.