Title

Energy Optimization For Wireless Video Transmission Employing Hybrid ARQ

Journal

IEEE Transactions on Vehicular Technology, 68(6)

ISSN

0018-9545

Authors

Zhang, Bentao
Cosman, Pamela
Milstein, Laurence B

Publication Date

2019

DOI

10.1109/tvt.2019.2907487

Peer reviewed

Energy Optimization For Wireless Video Transmission Employing Hybrid ARQ

Bentao Zhang, Pamela Cosman, *Fellow, IEEE*, and Laurence B. Milstein, *Fellow, IEEE*

***Abstract*—In this paper we investigate energy-optimized wireless video transmission employing a hybrid automatic repeat request. We formulate the problem as maximizing the video quality, subject to a constraint on the wireless transmission energy consumption. We consider multiple parameters in multiple layers in a wireless video transmission system: transmit power, alphabet size, FEC code rate, maximum number of transmissions, and unequal video data importance. An analytical framework is proposed to include these parameters, which allows us to divide this problem into two sub-problems: data transmission and unequal error protection. The problem is tackled by solving the two sub-problems, which are done by exhaustive search and convex optimization, respectively. Simulations of different videos show that the proposed scheme outperforms methods using conventional data transmission and/or unequal error protection.**

***Index Terms*—Energy optimization, wireless video transmission, hybrid ARQ, unequal error protection.**

I. Introduction

MOBILE wireless video traffic has been growing rapidly in recent years. According to a Cisco report, global mobile data traffic will increase sevenfold between 2016 and 2021 [1]. However, due to the unpredictable nature of the wireless channel, it is a challenging task to provide high quality video transmission. In addition, mobile devices usually have limited battery capacities [2]. Large energy consumption or transmit power increases the probability of successful transmission, and hence yields high video quality, but results in a short battery life. Therefore, it is an important task to improve wireless video quality with limited energy consumption.

In this paper, we consider a cross-layer optimization with a maximum number of retransmissions. We employ Hybrid Automatic Repeat reQuest (HARQ) [3]–[14], which is a combination of FEC and automatic repeat request (ARQ). Pure FEC may add unnecessary redundancy, whereas pure ARQ may require many retransmissions due to heavy losses for each single transmission. The authors in [15], [16] suggest that HARQ outperforms pure FEC and pure ARQ for wireless video transmission.

Manuscript received April 13, 2018; revised September 24, 2018 and February 21, 2019; accepted February 26, 2019. Date of publication March 26, 2019; date of current version June 18, 2019. This research was partially supported by the Center for Wireless Communications at UCSD and Army Research Office under Grant W911NF-14-1-0340. The review of this paper was coordinated by Dr. X. Ge. *(Corresponding author: Bentao Zhang.)*

The authors are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093 USA (e-mail: bez004@ucsd.edu; pcosman@ucsd.edu; lmilstein@ucsd.edu).

Digital Object Identifier 10.1109/TVT.2019.2907487

In this paper, we consider both unequal importance for the transmissions of the same video content, and unequal importance for different video contents, and we include the analysis of both HARQ and AMC. We consider multiple parameters in multiple layers in a wireless video transmission system: transmit power, alphabet size, FEC code rate, maximum number of transmissions and unequal frame importance. We decouple the problem into two sub-problems, data transmission and video UEP, which are solved by exhaustive search and convex optimization, respectively. Finally we conduct simulations to compare the performance of the proposed scheme to existing data HARQ and/or video UEP schemes. The results show that the proposed scheme significantly outperforms comparison HARQ and UEP methods.

The contributions of this work can be summarized as follows:

- We establish a framework to optimize video quality, subject to an energy consumption constraint, which considers both the unequal importance of multiple transmissions and the unequal importance of different video contents.
- We decouple a non-convex problem into two sub-problems, where the first one is non-convex and solved by exhaustive search, and the second one is convex and solved using a Lagrangian multiplier.
- We conduct simulations on different videos, and both simulation and theoretical results show that the proposed scheme significantly outperforms the baseline scheme which either takes into account only one of the unequal importances considered in this paper or does not consider any unequal importance at all.
- We analyze the performance of the algorithm using a specific set of parameters for all videos, instead of using the optimal parameters for each video. Results show that the specific set of parameters can provide close-to-optimal video quality, and thus the proposed algorithm can be run offline to generate a lookup table. During the video transmission, the parameters are obtained from the table, and the power and extra latency due to computation are avoided. The complexity of the offline algorithm is analyzed.

The rest of the paper is organized as follows. The related work is introduced in Section II. In Section III, we formulate the problem and propose an algorithm. In Sections IV and V, we introduce the two procedures in the algorithm. We show simulation results in Section VI, and conclusions are drawn in Section VII.

II. Related Work

A. Energy Optimization For Video Transmission

Energy optimization for video transmission has been studied by many researchers. In [17], the authors developed a framework to estimate the energy consumption of the video based on the channel state and video characteristics. However, the unequal importance of multiple transmissions was not fully investigated. In [18], the authors optimized the energy efficiency at the application layer. They adapt video encoder parameters based on the channel state and drop less important packets to avoid network congestion. In [19], a quasi-quadrature modulation is used to minimize the power consumption of video transmission. In [20], the authors proposed an algorithm to assign different forward error correction (FEC) code rates for the frames in a group of pictures (GOP) and drop packets with low priority. In [21], the authors studied coding unit level prioritization and FEC code rate selection in a wireless relay network. The authors in [22] studied the trade-off between energy saving and video quality and proposed an algorithm to assign different adaptive modulation and coding (AMC) modes to different video encoding layers. In [23], the authors maximized video quality, subject to an energy constraint in a D2D network, by optimizing the bit rate and FEC code rate on the paths between nodes in the D2D network. In [24], the optimal power allocation and the link adaptation algorithms are derived analytically, subject to a quality-of-service constraint. However, the fact that the video contents in a video sequence can require different quality-of-service, which reflects the unequal importance of video contents, is not taken into account. In [25], the source and channel coding are jointly optimized to achieve the best video quality, under an energy constraint, which is suitable for video streaming scenarios. In [26], the authors optimized energy efficiency for MIMO-OFDM multimedia communication with quality-of-service constraints. The trade-off of the energy between video encoding and transmission was discussed in [27].

B. Video Transmission With HARQ

We summarize the work on video transmission employing HARQ in this section. In [28], the authors proposed an APP/MAC/PHY cross-layer framework to optimize perceptual video quality. The maximum number of retransmissions varies with the layer in which the packet is encoded. If a transmission fails, the next retransmission is assigned a lower order modulation and coding scheme to provide a satisfactory probability of success. A limited-retransmission priority encoding transmission (LR-PET) scheme investigated PET with multiple transmissions [29]. The optimal protection depends on both the importance of the stream content and its behavior in future transmissions. The authors in [30] showed that under a scenario where the video contents have the same decoding deadline, the optimal strategy concludes that the most important packets should be retransmitted as often as needed and the less important ones may get discarded. A prioritized retransmission scheme was proposed based on the error propagation effect of the lost packet [31]. The authors in [32] assessed the impact of the lost macroblocks on the reconstructed frame and the ones with the highest impact are prioritized to be retransmitted. In [33], a video flow is subdivided into independent and incrementally encoded packets and the HARQ scheme privileges the retransmission of independent packets. The maximum number of transmissions is changed based on the importance of the packet for LTE networks in [34], [35]. The authors in [36] considered interlayer FEC and HARQ, and found the best FEC code rate distribution among the video layers to minimize the video distortion. In [37], retransmissions are used for the base layer to reduce network congestion. The authors in [38] studied unequal error protection (UEP), retransmission and GOP-level interleaving. They proposed segment-wise and byte-wise retransmission based on different types of receiver feedback. When the round trip time fluctuates significantly, the authors in [39] proposed a retransmission scheme to adaptively control the retransmission window size. The transmit power optimization problem was solved in [40] for multimedia applications, based on the observation that some performance metrics such as throughput, delay, and peak signal-to-noise ratio, may exhibit a staircase behavior for particular systems with HARQ. The authors in [41] proposed a scheduling scheme based on channel quality in LTE networks. In [42], the video source is able to adjust the video layers based on the channel estimator at the receiver. In [43], an adaptive unequal video protection method was designed for peer-to-peer video streaming over mobile wireless mesh networks, which assigns different priorities to the frames. A priority-based key frame protection method, which utilizes the idea of unequal error protection, was studied on a 5G network in [44]. A joint source-channel resource allocation problem was investigated in [45] to achieve adaptive error protection for video transmission.

The existing schemes on video retransmission employing HARQ can be categorized into two groups. For the first group, multiple transmissions of the same video content are treated equally, and the authors leverage the unequal importance of video contents which arises from the coding/decoding structure [46]. For example, for an IPPP encoding structure, a GOP begins with an intra-coded I frame followed by predicted P frames. An I frame does not use any other frames as reference for encoding, whereas a P frame uses the preceding frame as reference [47]. The frames at the beginning of a GOP have higher importance than later frames because of the decoding dependency of later frames on previous frames. Thus, different video contents should have unequal protection during transmission to maximize the decoded video quality. Examples of leveraging unequal importance of video contents include assigning different maximum numbers of transmissions, assigning different priorities for retransmission, and using different FEC codes. We analyze this unequal importance quantitatively in Section III-C, using the model in [48] and [49].

For the second group, the multiple transmissions of the same video content have unequal importance, but there lacks an analytical framework to provide the optimal strategy in each transmission. This unequal importance of multiple transmissions was studied using a single alphabet size and variable power [4],

[5]. The authors derived the optimal transmit power for each transmission round and this unequal power allocation outperforms the scheme in which the power is kept constant in each transmission. The optimal rate in each transmission for incremental redundancy was analyzed in [50], [51]. However, most of the existing data transmission HARQ schemes are based on information-theoretical analysis, which assumes there is no error when the SNR is larger than a threshold and assumes there is always an error when the SNR is smaller than the threshold. In our previous work [52], we used multiple alphabet sizes and variable power and investigated both the optimal transmit power and the optimal alphabet selection algorithm across the transmissions of the same data content, based on the actual packet error rate (PER) performance of a turbo code.

III. Problem Setup and Formulation

A. System Setup and Assumptions

A source encoder encodes a GOP, which contains T frames, with IPPP structure. Each frame is encoded into one or more slices. A slice is encoded by the FEC encoder and becomes a *codeword*, and is then sent through the wireless channel. A slice contains only information bits, a codeword is a slice plus parity bits, and a *packet* is any realization of a codeword. That is, a codeword can be transmitted multiple times, and the realization of each transmission is a packet. Slice copy is used at the decoder for error concealment. The term "packet error rate/probability" (PER) denotes the probability of error of any single transmission or retransmission of a codeword, and "overall packet error rate/probability" denotes the probability of error after a maximum number of transmissions of a codeword. For simplicity, the slices within frame i use the identical FEC code rate r_i, maximum number of transmissions N_i, power allocation algorithm and alphabet size mapping algorithm, where $i = 1, 2, ..., T$. The available FEC code rates are 1/2, 1/3 and 1/5, and the retransmissions of a codeword use the same FEC code rate as the first transmission. The maximum number of transmissions $N_i \in \{1, 2, ...N_{max}\}$. For power allocation, the transmit power varies with retransmission number of a codeword/slice, but does not vary with the slice number in a frame. Thus, in a given frame, the power depends on only the retransmission number of a codeword/slice, but not which slice it is. The maximum transmit power is S_{max}. For alphabet size mapping, adaptive M-PSK is used, and available modulations are BPSK, QPSK, 8PSK and 16PSK. Thus, retransmissions are allowed to use different alphabet sizes from the first transmission. The alphabet mapping algorithm will be explained in Section III-B. All packets contain L bits (including FEC, but excluding CRC bits and tail bits) and we ignore the influence of CRC bits and tail bits since their number is small relative to L. The symbol duration of the system, T_s, is kept constant, and a rectangular pulse shape is used. The number of symbols for an M-PSK packet is $L/log_2(M)$, and the number of information bits in a codeword in frame i is Lr_i. The receiver is a matched filter and coherent detection is used. The system block diagram is shown in Fig. 1.

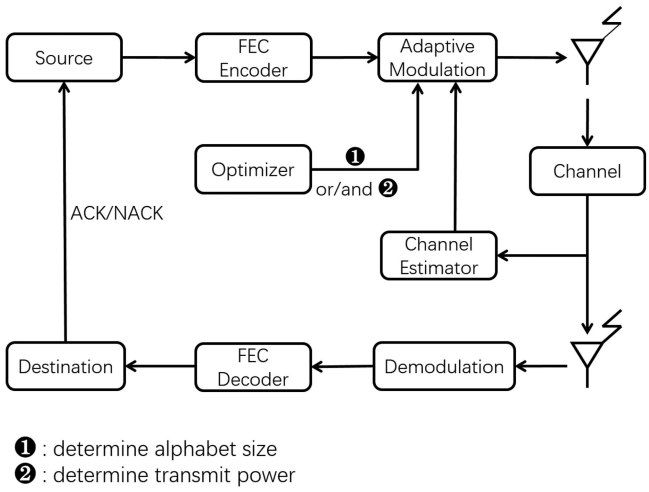


Fig. 1. System block diagram.

The instantaneous channel gain for the j-th transmission of a codeword in frame i, $\gamma_{j,i}$, is assumed to be Rayleigh distributed, and the pdf of $\gamma_{j,i}$ does not change with time. The fades of different transmissions of a codeword are assumed to be independent, and we assume perfect CSI is available at the transmitter.

B. Alphabet Size Mapping

Since adaptive modulation is used, we introduce the alphabet size mapping in this subsection before we formulate the problem. The instantaneous received signal-to-noise ratio (SNR) for the j-th transmission of a codeword in frame i is $\Gamma_{j,i} = \gamma_{j,i}^2 S_{j,i} T_s / N_0$, where $S_{j,i}$ is the transmit power for the j-th transmission of a codeword in frame i and N_0 is the spectral density of the additive Gaussian noise. Since the packet error probability for turbo codes cannot be expressed analytically, we fit the packet error probability with an exponential function as in [3]. For a given code rate, we have

$$\psi_M(\Gamma_{j,i}) = \begin{cases} 1, & 0 < \Gamma_{j,i} < \Gamma_M^{min} \\ a_{log_2 M} e^{-b_{log_2 M} \Gamma_{j,i}}, & \Gamma_{j,i} \geq \Gamma_M^{min}, \end{cases} \quad (1)$$

where $\psi_M(\Gamma_{j,i})$ is the conditional single-transmission packet error probability, conditioned on $\Gamma_{j,i}$, for M-PSK, and $a_n > 0, b_n > 0, n = 1, 2, 3, 4$. The parameters a_n, b_n and Γ_M^{min} are obtained through simulation and curve fitting and they depend on the code rate. The simulated packet error rate for a rate 1/3 turbo code is shown in Fig. 2.

We map alphabet sizes as in [3]. We set a PER "upper bound" $U_{j,i}$ for the j-th transmission of the codewords in the i-th frame. $U_{j,i}$ and $U_{k,i}$ can be the same or different for $j \neq k$. Let the instantaneous received SNR boundaries for the j-th transmission of the slices in the i-th frame be $(\Gamma_{j,i}^{(1)}, \Gamma_{j,i}^{(2)}, \Gamma_{j,i}^{(3)}, \Gamma_{j,i}^{(4)}, \Gamma_{j,i}^{(5)})$, and choose M-PSK when $\Gamma_{j,i}^{(log_2 M)} \leq \Gamma_{j,i} < \Gamma_{j,i}^{(log_2 M+1)}$. The SNR boundaries for the j-th transmission of the codewords in the i-th frame are determined by setting the left hand side of

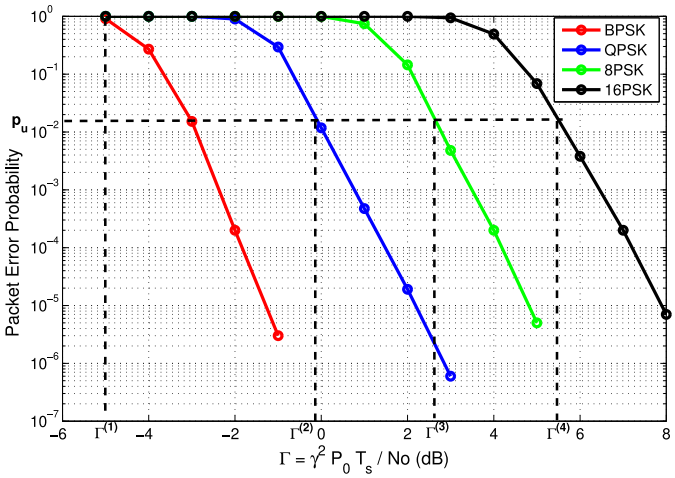


Fig. 2. Packet error rate vs. instantaneous received SNR.

Equation (1) equal to $U_{j,i}$, where $\Gamma_{j,i}^{(1)} = \Gamma_2^{min}$, $\Gamma_{j,i}^{(5)} = \infty$:

$$\Gamma_{j,i}^{(n)} = \frac{1}{b_n} ln\left(\frac{a_n}{U_{j,i}}\right), \quad n = 2, 3, 4. \tag{2}$$

The terms $\Gamma_{j,i}^{(2)}, \Gamma_{j,i}^{(3)}$ and $\Gamma_{j,i}^{(4)}$ are determined by $U_{j,i}$. Thus, we aim to choose the largest alphabet size so that the curve-fitted conditional PER, conditioned on $\Gamma_{j,i}$, is lower than $U_{j,i}$ when $\Gamma_{j,i} > \Gamma_{j,i}^{(2)}$, and we choose BPSK when $\Gamma_{j,i}^{(1)} < \Gamma_{j,i} \leq \Gamma_{j,i}^{(2)}$, i.e., the curve fit to the conditional PER is lower than 1. The system does not transmit when $\Gamma_{j,i} \leq \Gamma_{j,i}^{(1)}$, i.e., the fitted conditional PER is 1. Although $U_{j,i}$ is not a strict upper bound on the conditional PER over the whole range where BPSK is used, we use the term "upper bound" since the conditional PER is lower than $U_{j,i}$ for all of the channel states which are above some threshold.

C. Problem Formulation

We want to minimize the mean square error (MSE) in the GOP of the given video sequence, subject to overall energy and maximum power constraints. The system parameters we can directly control are the maximum number of transmissions, FEC code rate, power, and PER "upper bound". According to [48] and [49], the MSE directly depends on the average PERs for the frames (although the average PER further depends on the system parameters above). Here the "average PER" is the instantaneous PER averaged over the Rayleigh fading channel. If the average PER for frame i in a GOP is denoted by P_{ei}, where $i = 1, 2, ..., T$, then the MSE incurred in frame i is

$$\sigma_0^2 \sum_{\tau=0}^{i} P_{e\tau} \frac{1}{1 + \alpha(i - \tau)}, \tag{3}$$

where σ_0^2 are α are parameters which depend on the video. The above model assumes that 1) the pixel errors incurred in different frames are independent and 2) the pixel errors incurred in a frame, but propagated from different frames, are uncorrelated. The first assumption is valid in our model, since a packet can only contain contents from one frame, and different packets experience independent fading, thus the pixel errors in different frames are independent. The second assumption is realistic when P_{ei} is small. The accuracy of the model is evaluated in Section V. Therefore, the MSE over the GOP is

$$\begin{aligned} MSE(\{P_{ei}\}) &= \sigma_0^2 \sum_{i=1}^{T} \sum_{\tau=1}^{i} P_{e\tau} \frac{1}{1 + \alpha(i - \tau)} \\ &= \sigma_0^2 \sum_{i=1}^{T} P_{ei} \sum_{\tau=0}^{T-i} \frac{1}{1 + \alpha\tau}, \end{aligned} \tag{4}$$

where $\{P_{ei}\} = (P_{e1}, P_{e2}, ..., P_{eT})$ is the set of PERs for T frames. Note that the MSE over the GOP in Equation (4) is just the sum of the MSE of every frame in the GOP, so Equation (4) does not require further assumptions beyond those made for Equation (3).

The optimization problem is formulated as follows:

$$\begin{aligned} min \quad & MSE(\{P_{ei}\}) \\ s.t. \quad & \sum_{i=1}^{T} n_i \mathcal{E}(P_{ei}) = min(\mathcal{E}_c, \tilde{\mathcal{E}}_c) \\ & 0 \leq S_{j,i} \leq S_{max} \quad \text{for } i = 1, 2, ..., T \text{ and} \\ & \qquad\qquad\qquad\qquad j = 1, 2, ..., N_i \\ variables: \quad & \{P_{ei}\} \end{aligned} \tag{5}$$

where n_i is the number of information bits in frame i, $\mathcal{E}_c$ is the energy constraint for the entire GOP, $\tilde{\mathcal{E}}_c = (\sum_{i=1}^{T} n_i) S_{max} T_s N_{max} / r_{min}$ represents the maximum energy that a GOP could consume. Here, r_{min} is the lowest FEC code rate (r_{min} = 1/5 in this paper). The maximum energy occurs when the FEC is the strongest possible and the maximum number of transmissions is used. The term $\mathcal{E}(P_{ei})$ is the energy consumption per information bit in frame i to achieve an average PER of P_{ei}, i.e., the energy consumption of a codeword divided by the number of information bits in the codeword. The units of $\mathcal{E}_c$, $\tilde{\mathcal{E}}_c$ and $\mathcal{E}(P_{ei})$ are Joules. We define $\mathcal{E}(P_{ei})$ in this way because the number of information bits in frame i is determined by the video encoder, which does not depend on the FEC rate r_i. Thus, $\mathcal{E}(P_{ei})$ is a fair comparison for different FEC rates. The set operator is $\{\cdot\}$. The maximum energy that an information bit in frame i can possibly use is $S_{max} T_s N_{max} / r_{min}$, i.e., the codeword uses BPSK, FEC rate r_{min}, maximum power S_{max} and gets transmitted N_{max} times. The energy constraint $\mathcal{E}_c$ is used to minimize MSE when $\mathcal{E}_c \leq \tilde{\mathcal{E}}_c$, and the energy consumption cannot exceed $\tilde{\mathcal{E}}_c$ when $\mathcal{E}_c > \tilde{\mathcal{E}}_c$. In Equation (5), the function $\mathcal{E}(P_{ei})$ is not uniquely defined in the above formulation since there can be multiple sets of parameters $(N_i, r_i, S_{j,i}, U_{j,i})$ which generate the same P_{ei}, but do not result in the same energy consumption. Thus, for a given P_{ei}, there can be multiple values of energy $\mathcal{E}(P_{ei})$.

As a consequence, for any P_{ei} being considered, we choose the energy-minimizing $\mathcal{E}_{min}(P_{ei})$ to achieve P_{ei}, where $\mathcal{E}_{min}(P_{ei})$ is defined as the least energy consumption per

TABLE I
TABLE OF VARIABLES

$\{\cdot\}$	the set of a variable
T_s	symbol duration
N_0	spectral density of Gaussian noise
L	number of coded bits in a packet
T	number of frames in a GOP
n_i	number of slices/codewords in frame i
N_i	maximum number of transmissions for a codeword in frame i
r_i	code rate for a codeword in frame i
$\gamma_{j,i}$	instantaneous channel gain for the j-th transmission of a codeword in frame i
$\Gamma_{j,i}$	instantaneous received SNR for the j-th transmission of a codeword in frame i
$f_{\Gamma_{j,i}}(\Gamma_{j,i})$	pdf of $\Gamma_{j,i}$ for the j-th transmission of a codeword in frame i
$S_{j,i}$	power for the j-th transmission of a codeword in frame i
$U_{j,i}$	upper bound on conditional PER for the j-th transmission of a codeword in frame i
$\Gamma^{(j)}$	SNR boundary for alphabet size
$\psi_M(\Gamma)$	conditional PER for M-PSK using turbo code
P_{ei}	average overall PER for a codeword in frame i
$MSE(\{N_i\},\{r_i\},\{S_{j,i}\},\{U_{j,i}\})$	function of the system parameters to calculate the MSE over GOP
$MSE(\{P_{ei}\})$	function of the average frame PER to calculate the MSE over GOP
$\mathcal{E}^{ave,N_i,r_i}(S_{1,i},...,S_{N_i,i},U_{1,i},...,U_{N_i,i})$	function to calculate the average overall energy for a codeword in frame i
$P^{ave,N_i,r_i}(S_{1,i},...,S_{N_i,i},U_{1,i},...,U_{N_i,i})$	function to calculate the average overall PER for a codeword in frame i
$\mathcal{E}_{min}^{N_i,r_i}(P_{ei})$	minimum average energy to achieve an average overall PER of P_{ei}, with given N_i and r_i

information bit to achieve an average PER of P_{ei}. Notice that $\mathcal{E}_{min}(P_{ei})$ is a function of P_{ei}, but not the system parameters $(N_i, r_i, S_{j,i}, U_{j,i})$. Intuitively, if a particular packet is to be transmitted with a certain PER, it does not make sense to achieve that PER using a parameter set $(N_i, r_i, S_{j,i}, U_{j,i})$ that uses a higher energy if the same PER can be achieved using a parameter set that uses a lower energy. Hence in the following, as the algorithm searches through the possible values of PER, for each value of PER, only the energy-minimizing value of achieving that PER, $\mathcal{E}_{min}(P_{ei})$, will be considered.

Note that there are two constraints in Equation (5): energy constraint $\mathcal{E}_c$ and power constraint S_{max}. The energy $\mathcal{E}_c$ constraint is determined by the video application, and is determined by such factors as the desired video quality level and the battery status. The power constraint S_{max} is determined by the hardware, e.g., the power amplifier. These two constraints are determined independently.

D. Algorithm

Equation (5) is a non-convex optimization problem due to the non-convexity of the term $\mathcal{E}(P_{ei})$, so we solve it by using a combination of exhaustive search and Lagrangian multiplier method. First, we quantize P_{ei}. For each quantized value of P_{ei}, we then quantize the variables $S_{1,i},...,S_{N_i,i},U_{1,i},...,U_{N_i,i}$ and exhaustively search all possible tuples of $N_i, r_i, S_{1,i},...,S_{N_i,i},U_{1,i},...,U_{N_i,i}$ to find all the tuples which yield PER P_{ei}. Among these, we find the one which has the minimum energy consumption $\mathcal{E}_{min}(P_{ei})$. After examining all quantized values of P_{ei} and obtaining the corresponding minimum energy, we get the function $\mathcal{E}_{min}(P_{ei})$. Then we use the Lagrangian multiplier method to solve Equation (5).

Before introducing the details of the algorithm, we define $\mathcal{E}^{ave}(N_i, r_i, S_{1,i},...,S_{N_i,i},U_{1,i},...,U_{N_i,i})$ to be the average overall energy consumption (including retransmissions) of an information bit in frame i, i.e., the overall energy consumption of a codeword divided by the number of information bits in the codeword, and we define $P^{ave}(N_i, r_i, S_{1,i},...,S_{N_i,i},U_{1,i},...,U_{N_i,i})$ to be the average overall PER for a codeword in frame i. Notice that the codewords in one frame have the same $N_i, r_i, S_{j,i}, U_{j,i}$, but the modulation alphabet is allowed to vary from one codeword to the next based on the channel state.

IV. PROPOSED HARQ

A. Problem Formulation

The procedure which obtains $\mathcal{E}_{min}(P_{ei})$ solves the HARQ problem by optimizing the parameters in each transmission, and does not depend on frame number i, so we simplify the notation by removing the dependence on i. Then the problem can be formulated as follows:

$$\begin{aligned} min \quad & \mathcal{E}^{ave}(N, r, \{S\}, \{U\}) \\ s.t. \quad & P^{ave}(N, r, \{S\}, \{U\}) = P_e \\ & 0 \le S_j \le S_{max} \quad j = 1, 2, ..., N \\ variables: \quad & N, r, \{S\}, \{U\}, \end{aligned} \tag{6}$$

where N is the maximum number of transmissions of the codeword, r is the FEC code rate of the codeword, $\{S\} = (S_1, ..., S_N)$ with S_j being the transmit power for the j-th transmission of the codeword and $\{U\} = (U_1, ..., U_N)$ with U_j being the PER "upper bound" for the j-th transmission of the codeword. $\mathcal{E}^{ave}(N, r, \{S\}, \{U\})$ is the average overall energy consumption (including retransmissions) of an information bit in a codeword, $P^{ave}(N, r, \{S\}, \{U\})$ is the average overall PER for a codeword, and P_e is the average overall PER constraint. We define $\{\Gamma\} = (\Gamma_1, \Gamma_2, ..., \Gamma_M)$ as the set of channel states for M transmissions of the codeword, where $M \le N$, and $\{X\} = (X_1, X_2, ..., X_M)$ as the set of outcomes for M transmissions of the codeword. $X_j = 1$ means the j-th transmission

of the codeword is successful, and $X_j = 0$ means the j-th transmission is not successful. Let $P_e^{cndt}(N, r, \{S\}, \{U\} \mid \{\Gamma\})$ be the conditional codeword error probability, conditioned on $\{\Gamma\}$, i.e., the probability that the codeword cannot be successfully decoded after N transmissions with the given channel states $\{\Gamma\}$ in the transmissions. Let $\mathcal{E}^{cndt}(N, r, \{S\}, \{U\} \mid \{\Gamma\}, \{X\})$ be the conditional energy consumption of an information bit, conditioned on $\{\Gamma\}$ and $\{X\}$, i.e., the sum of the energy consumptions of all transmissions of the codeword with the given channel states $\{\Gamma\}$, divided by the number of information bits in the codeword. Then we have the following equations:

$$P_e^{cndt}(N, r, \{S\}, \{U\} \mid \{\Gamma\}) = \mathcal{P}(X_1 = 0, X_2 = 0, ..., X_N = 0 \mid \{\Gamma\}) \quad (7)$$

$$\mathcal{E}^{cndt}(N, r, \{S\}, \{U\} \mid \{\Gamma\}, \{X\}) = \begin{cases} \sum_{j=1}^{k} e_j(\Gamma_j) & \text{if } k \text{ is the smallest number such that} \\ & X_j = 1, k \le N \\ \sum_{j=1}^{N} e_j(\Gamma_j) & \text{if } X_j = 0 \text{ for all } j \in [1, N-1] \end{cases}, \quad (8)$$

where $e_j(\Gamma_j)$ is the energy consumption during the j-th transmission of an information bit, given that the channel state is Γ_j.

B. Special Case

For simplicity of presentation, consider the special case where $N = 2$. Other cases can be shown to follow by similar analysis. The average overall PER is

$$\begin{aligned} P^{ave}(2, r, \{S\}, \{U\}) &= E_{\{\Gamma\}}[P_e^{cndt}(2, r, \{S\}, \{U\} \mid \{\Gamma\})] \\ &= \int_0^\infty \int_0^\infty P_e^{cndt}(2, r, \{S\}, \{U\} \mid \{\Gamma\}) f_{\Gamma_1,\Gamma_2}(\Gamma_1, \Gamma_2) d\Gamma_1 d\Gamma_2 \\ &= \int_0^\infty \int_0^\infty P_{e1}^{cndt}(r, S_1, U_1 \mid \Gamma_1) P_{e2}^{cndt}(r, S_2, U_2 \mid \Gamma_2) \\ &\quad \times f_{\Gamma_1,\Gamma_2}(\Gamma_1, \Gamma_2) d\Gamma_1 d\Gamma_2 \\ &= \int_0^\infty P_{e1}^{cndt}(r, S_1, U_1 \mid \Gamma_1) f_{\Gamma_1}(\Gamma_1) d\Gamma_1 \\ &\quad \times \int_0^\infty P_{e2}^{cndt}(r, S_2, U_2 \mid \Gamma_2) f_{\Gamma_2}(\Gamma_2) d\Gamma_2, \end{aligned} \quad (9)$$

where $f_{\Gamma_1,\Gamma_2}(\Gamma_1, \Gamma_2)$ is the joint pdf of the channel states for the two transmissions, $f_{\Gamma_j}(\Gamma_j)$ is the pdf of the channel state for the j-th transmission, and $P_{ej}^{cndt}(r, S_j, U_j \mid \Gamma_j)$ is the conditional PER in the j-th transmission, conditioned on channel state Γ_j. The SNR boundaries in the j-th transmission $(\Gamma_j^{(1)}, \Gamma_j^{(2)}, \Gamma_j^{(3)}, \Gamma_j^{(4)}, \Gamma_j^{(5)})$ are obtained from the method in Section III-B. With the SNR boundaries, we have

$$P_{ej}^{cndt}(r, S_j, U_j \mid \Gamma_j) = \psi_{2^i}(\Gamma_j) = a_i e^{-b_i \Gamma_j} \quad \text{when } \Gamma_j^{(i)} < \Gamma_j < \Gamma_j^{(i+1)}, i = 1, 2, 3, 4 \quad (10)$$

where the curve-fitted parameters a_i and b_i depend on r. Let A_j be the average SNR in the j-th transmission, where $A_j = E[\gamma^2 S_j T_s / N_0]$. Then we have

$$\begin{aligned} &P^{ave}(N, r, \{S\}, \{U\}) \\ &= \frac{1}{A_1 A_2} \left[\int_0^{\Gamma_1^{(1)}} e^{-\frac{\Gamma_1}{A_1}} d\Gamma_1 + \sum_{i=1}^{4} a_i \int_{\Gamma_1^{(i)}}^{\Gamma_1^{(i+1)}} e^{-\left(b_i + \frac{1}{A_1}\right)\Gamma_1} d\Gamma_1 \right] \\ &\quad \times \left[\int_0^{\Gamma_2^{(1)}} e^{-\frac{\Gamma_2}{A_2}} d\Gamma_2 + \sum_{i=1}^{4} a_i \int_{\Gamma_2^{(i)}}^{\Gamma_2^{(i+1)}} e^{-\left(b_i + \frac{1}{A_2}\right)\Gamma_2} d\Gamma_2 \right] \end{aligned} \quad (11)$$

The average overall energy per information bit is in Equation (12), where $\mathcal{E}_j^{cndt}(r, S_j, U_j \mid \Gamma_j)$ is the conditional energy consumption during the j-th transmission of the information bit, conditioned on channel state Γ_j.

$$\begin{aligned} &\mathcal{E}^{ave}(2, r, \{S\}, \{U\}) \\ &= E_{\{\Gamma\},\{X\}}[\mathcal{E}^{cndt}(2, r, \{S\}, \{U\} \mid \{\Gamma\}, \{X\})] \\ &= E_{\{\Gamma\}}[\mathcal{P}(X_1 = 1)\mathcal{E}^{cndt}(2, r, \{S\}, \{U\} \mid \{\Gamma\}, X_1 = 1) \\ &\quad + \mathcal{P}(X_1 = 0)\mathcal{E}^{cndt}(2, r, \{S\}, \{U\} \mid \{\Gamma\}, X_1 = 0)] \\ &= E_{\{\Gamma\}}\big[\mathcal{P}(X_1 = 1)\mathcal{E}_1^{cndt}(r, S_1, U_1 \mid \Gamma_1) + \mathcal{P}(X_1 = 0) \\ &\quad \times \left(\mathcal{E}_1^{cndt}(r, S_1, U_1 \mid \Gamma_1) + \mathcal{E}_2^{cndt}(r, S_2, U_2 \mid \Gamma_2)\right)\big] \\ &= \int_0^\infty \int_0^\infty \big[\mathcal{E}_1^{cndt}(r, P_1, U_1 \mid \Gamma_1) + P_{e1}^{cndt}(r, P_1, U_1 \mid \Gamma_1) \\ &\quad \times \mathcal{E}_2^{cndt}(r, S_2, U_2 \mid \Gamma_2)\big] f_{\Gamma_1,\Gamma_2}(\Gamma_1, \Gamma_2) d\Gamma_1 d\Gamma_2 \\ &= \int_0^\infty \mathcal{E}_1^{cndt}(r, S_1, U_1 \mid \Gamma_1) f_{\Gamma_1}(\Gamma_1) d\Gamma_1 \\ &\quad + \int_0^\infty P_{e1}^{cndt}(r, S_1, U_1 \mid \Gamma_1) f_{\Gamma_1}(\Gamma_1) d\Gamma_1 \\ &\quad \times \int_0^\infty \mathcal{E}_2^{cndt}(r, S_2, U_2 \mid \Gamma_2) f_{\Gamma_2}(\Gamma_2) d\Gamma_2 \\ &= \frac{1}{A_1 L r} \sum_{i=1}^{4} \int_{\Gamma_1^{(i)}}^{\Gamma_1^{(i+1)}} \frac{S_1 T_s}{i} e^{-\frac{\Gamma_1}{A_1}} d\Gamma_1 + \frac{1}{A_1 A_2 L r} \\ &\quad \times \left[\int_0^{\Gamma_1^{(1)}} e^{-\frac{\Gamma_1}{A_1}} d\Gamma_1 + \sum_{i=1}^{4} a_i \int_{\Gamma_1^{(i)}}^{\Gamma_1^{(i+1)}} e^{-(b_i + \frac{1}{A_1})\Gamma_1} d\Gamma_1 \right] \\ &\quad \times \left[\sum_{i=1}^{4} \int_{\Gamma_2^{(i)}}^{\Gamma_2^{(i+1)}} \frac{S_2 T_s}{i} e^{-\frac{\Gamma_2}{A_2}} d\Gamma_2 \right], \end{aligned} \quad (12)$$

With the above expressions, we would like to solve Equation (6). However, the problem is non-convex, even when $N = 2$. We use an exhaustive search to find the optimal solution. First, we quantize P_e. For each quantized value of P_e, we then quantize the variables $S_1, ..., S_N, U_1, ..., U_N$ and exhaustively search all possible tuples of $N, r, S_1, ..., S_N, U_1, ..., U_N$ to find the one which gives average PER P_e and has the minimum energy con-

TABLE II
SIMULATION PARAMETERS

T	30
number of information bits in a packet	256
T_s	10^{-6}s
S_0	5mW
N_0	3.98×10^{-21}W/Hz
N	3
r	1/5
video size	1024x576
number of frames in video	1903 to 2229
$\mathcal{E}_c$	20mJ or 10mJ

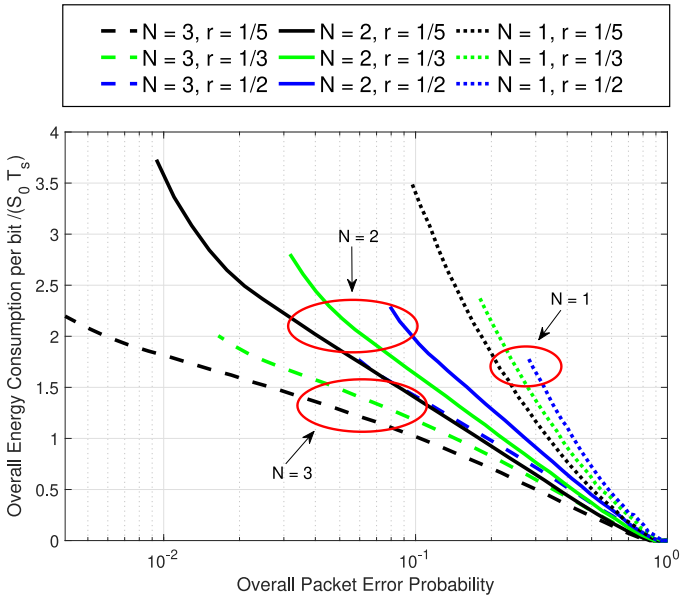


Fig. 3. Energy consumption for different FEC rates and maximum number of transmissions. Average channel SNR is 0 dB.

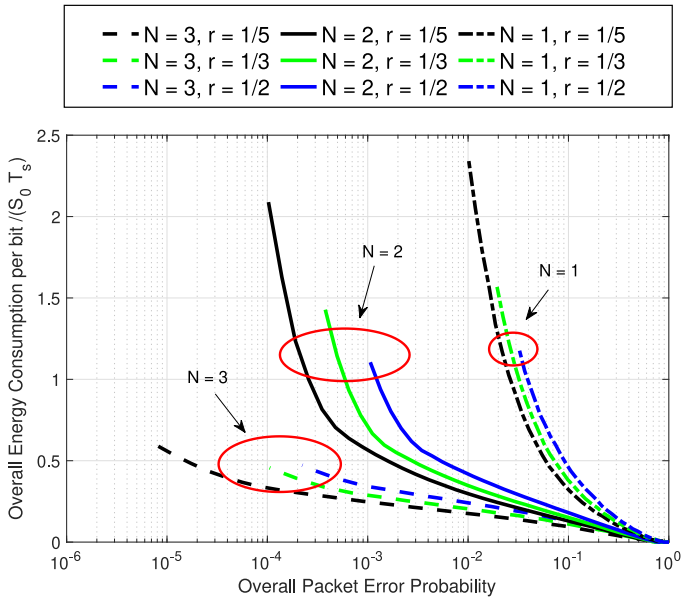


Fig. 4. Energy consumption for different FEC rates and maximum number of transmissions. Average channel SNR is 10 dB.

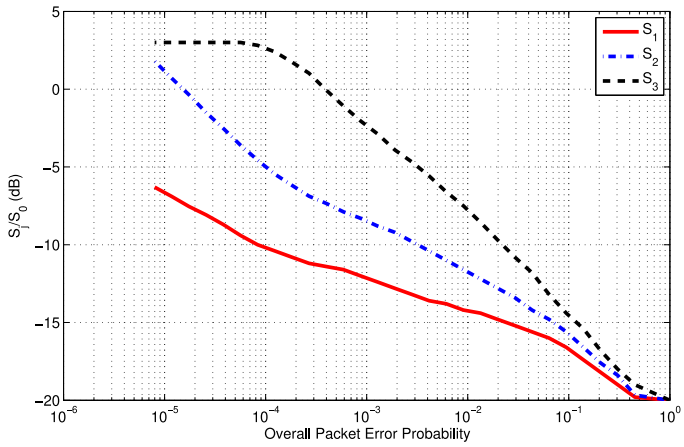


Fig. 5. Power distribution for $N = 3$ and $r = 1/5$. Average channel SNR is 10 dB.

sumption $\mathcal{E}_{min}(P_e)$. We quantize three continuous parameters: P_e, S and U. In the numerical exhaustive search, the step size for $log_{10}(P_e)$ is 0.02, the step size for $10log_{10}(S)$ is 0.02 and the step size for $log_{10}(U)$ is 0.02. The range for P_e is from the minimum achievable PER to 1. The minimum achievable PER depends on the channel statistics, e.g., 4×10^{-3} when $E[\gamma^2 S_0 T_s/N_0] = 1$. The range for S is from 0 to $S_{max} = 2S_0$. The range for U is from 10^{-6} to 1.

Figs. 3 and 4 show the minimum energy consumption per bit vs. PER for different (N, r) combinations when $E[\gamma^2 S_0 T_s/N_0]$ = 0 dB and 10 dB, respectively. The energy consumption per bit is normalized by $S_0 T_s$, i.e., the energy consumption per bit (in units of Joules) for BPSK without FEC in a single transmission. Thus, the y-axis is dimensionless. In this paper, we let the available protection combinations be $\{N\} = \{3, 2, 1\}, \{r\} = \{1/5, 1/3, 1/2\}$, so there are 9 possible protections. In both figures, each line color corresponds to a value of N: black for N = 3, green for $N = 2$, and blue for $N = 1$, and each line type corresponds to a value of r: dashed for $r = 1/5$, solid for $r = 1/3$, and dotted for $r = 1/2$. Then $\mathcal{E}_{min}(P_e)$ is the minimum among these curves. From both figures, we find that, given N, energy decreases with the decreasing of r, which is due to the coding gain; given r, energy decreases with the increasing of N, which is due to the diversity in the transmissions. The influence of N is larger than the influence of r, especially when the channel SNR is large. No matter what the channel SNR is, $(N = 3, r = 1/5)$ results in the minimum energy due to the best coding gain and diversity in transmissions.

Figs. 5 and 6 show the power and PER "upper bound" distributions for $(N = 3, r = 1/5)$ when $E[\gamma^2] = 10$ dB. In Fig. 5, the y-axis is the ratio of the transmit power to the constant power S_0 in dB. In Fig. 6, the y-axis is the parameter PER "upper bound", which determines the alphabet size based on the instantaneous CSI. The first transmission has the least power and the largest PER "upper bound" among the three transmissions, and the opposite holds for the third transmission. The conclusion is if there are three transmission opportunities, the first transmission should be sent at a low cost, because if it happens to succeed,

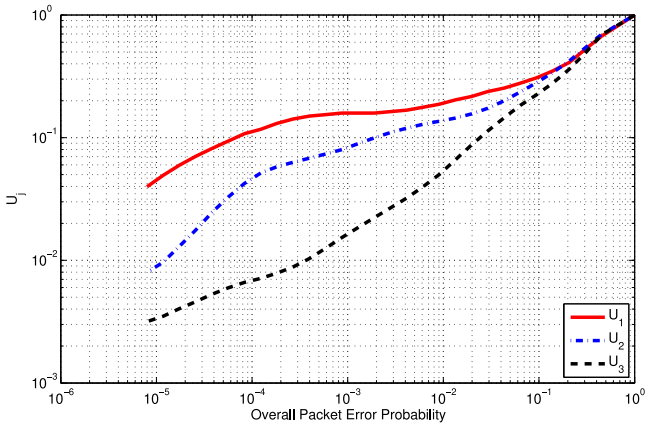


Fig. 6. PER "upper bound" distribution for $N = 3$ and $r = 1/5$. Average channel SNR is 10 dB.

then the second and third transmissions are not necessary and energy consumption is low; however, if the first transmission fails, we need to expend additional power and transmission duration on the second transmission to guarantee the PER; if the second transmission also fails, the third transmission should use the most resources since it is the last transmission opportunity and its transmission failure would result in the failure of the codeword.

C. Complexity Analysis

Since the algorithm uses an exhaustive search, the complexity of the algorithm is $O(\sum_{n=1}^{N_{max}} |\{r\}| \cdot (M_S M_U)^n) = O(|\{r\}| \cdot \frac{(M_S M_U)^{N_{max}+1}-1}{M_S M_U - 1})$, where $|\{r\}|$ is the number of rate options, and M_S and M_U are the number of discrete values of S and U in the exhaustive search, respectively. This number is exponentially increasing with N_{max}. Note, however, this optimization does not depend on the video and can be done offline. The results for the optimization, i.e., $N, r, \{S\}, \{U\}$, can be stored, and we only need a lookup table during the video transmission.

V. Proposed UEP

A. Problem Formulation

After obtaining $\mathcal{E}_{min}(P_{ei})$, we solve the video content UEP problem, since the PERs for each frame are optimized, and thus different frames have unequal error protection. This procedure can be formulated as the following problem:

$$\begin{aligned} min \quad & MSE(\{P_{ei}\}) \\ s.t. \quad & \sum_{i=1}^{T} n_i \mathcal{E}_{min}(P_{ei}) = min(\mathcal{E}_c, \tilde{\mathcal{E}}_c) \\ & 0 \le S_{j,i} \le S_{max} \quad \text{for } i = 1, 2, ..., T \text{ and} \\ & \qquad\qquad\qquad\qquad j = 1, 2, ..., N_i \\ variables: \quad & \{P_{ei}\}. \end{aligned} \tag{13}$$

Let the average overall PERs for the frames be $P_{e1}, P_{e2}, ..., P_{eT}$. The task is to find the optimal average PER for each frame, where frame i uses energy $\mathcal{E}_{min}(P_{ei})$ to achieve this PER. It is seen in Equation (3) that the error in the previous frames can propagate to later frames, so the frames at the front should have lower PER than frames at the end, i.e., the PER sequence should be in ascending order, i.e., $P_{ej} \le P_{ek}$ where $j \le k$.

Section IV showed that all frames should use the strongest protection, i.e, $N = 3$ and $r = 1/5$, to minimize the energy consumption. For simplicity of notation, let $g(P_{ei}) = \mathcal{E}_{min}(P_{ei})$. As stated in Section III-C, we know the solution to Equation (5) when $\mathcal{E}_c \ge \tilde{\mathcal{E}}_c$: the codeword uses BPSK, FEC rate r_{min}, maximum power S_{max}, and transmits N_{max} times. So we consider the case $\mathcal{E}_c < \tilde{\mathcal{E}}_c$. Then using Equation (4), Equation (13) can be written as

$$\begin{aligned} min \quad & \sigma_0^2 \sum_{i=1}^{T} P_{ei} \sum_{\tau=0}^{T-i} \frac{1}{1+\alpha\tau} \\ s.t. \quad & \sum_{i=1}^{T} n_i g(P_{ei}) = \mathcal{E}_c. \end{aligned} \tag{14}$$

where the constraint on $S_{j,i}$ has been taken into account in the procedure that yields $\mathcal{E}_{min}(P_{ei})$.

We obtained the numerical results of $g(P_{ei})$ through exhaustive search. We approximate the second order derivative by $\frac{g(P_{ei}+h)-2g(P_{ei})+g(P_{ei}-h)}{h^2}$, where h is a sufficiently small value. We find that the approximate second derivative is positive, so $g(P_{ei})$ is convex. Since the objective function in Equation (14) is a linear function of P_{ei}, it is a convex function; since $g(P_{ei})$ is convex, the constraint is convex, so this problem is convex and we can use the Lagrange multiplier method to get the globally optimal solution. The Lagrange function is

$$L = \sigma_0^2 \sum_{i=1}^{T} P_{ei} \sum_{\tau=0}^{T-i} \frac{1}{1+\alpha\tau} + \lambda \left(\sum_{i=1}^{T} n_i g(P_{ei}) - \mathcal{E}_c \right) \tag{15}$$

Letting $\frac{\partial L}{\partial P_{ei}} = 0$ for $i = 1, 2, ..., T$, we have

$$\sigma_0^2 P_{ei} \sum_{\tau=0}^{T-i} \frac{1}{1+\alpha\tau} + \lambda n_i g'(P_{ei}) = 0, \tag{16}$$

where $g'(P_{ei}) \triangleq \frac{g(P_{ei}+h)-g(P_{ei})}{h}$. Letting $\frac{\partial L}{\partial \lambda} = 0$, we have

$$n_i g(P_{ei}) - \mathcal{E}_c = 0. \tag{17}$$

Equations (16) and (17) together give the solution to Equation (14).

B. Complexity Analysis

The complexity of the proposed UEP algorithm is composed of two parts: parameter estimation of σ_0^2 and α, and solving Equations (16) and (17). The complexity for the first part depends on the algorithm of curve fitting and the complexity for the second part depends on the algorithm of solving the convex problem. This work does not focus on the complexity analysis of these two parts. However, as shown in Section VI-C, the UEP

algorithm can use fixed parameters without much loss of accuracy and performance and thus can be solved offline. Therefore, the results can be stored, and a lookup table is sufficient during video transmission.

VI. Numerical Results

A. Comparison Algorithms

The comparison HARQ scheme [4], [5] uses a single alphabet size and variable power, whereas the proposed HARQ uses variable alphabet sizes and variable power. The comparison UEP is similar to the method in [46], but with a modification that improves its performance. In [46], the authors assume constant power and a single alphabet size without retransmission, and they use three rates of FEC for the frames in a GOP. The video quality depends on the choice of these three values. The number of FEC code rates is limited due to the implementation complexity. We improve upon their method by allowing variable power and variable alphabet size and using the strongest FEC and the maximum number of transmissions for all frames, and assign three levels of energy consumption for the frames. As discussed in Section IV, using variable power with a single strong FEC performs better than using constant power and three different FEC values as done in [46]. Since the optimal energy levels are hard to obtain, we determine three non-overlapping ranges by trial and error, and randomly choose the three energy values from these non-overlapping ranges. The best ranges we found are $(1.2E_0, 1.5E_0]$, $(0.8E_0, 1.2E_0]$ and $(0.5E_0, 0.8E_0]$, where E_0 is a parameter dependent on the energy constraint. Since the video quality depends on the selection of these three values, we repeat this experiment 20 times and show the average Y-PSNR, where Y-PSNR is the peak signal-to-noise ratio (PSNR) for the Y channel in the YUV video. Here, PSNR $= 10log_{10}(V_{max}^2/MSE)$, V_{max} is the maximum pixel value, and MSE is the mean square error on the Y channel.

B. Simulation Environment

We use JM 15.0 software to achieve H.264 video coding. The previous frame is used for error concealment. We use video from a mobile video database [53], [54]. The video scenes are a football game with high motion, a person skating with medium motion, and an interview scene with slow motion. The videos are of size 1024×576 with 1903 frames, 2022 frames and 2229 frames. The video frame rate is 30 fps, so their lengths are approximately 60 to 70 seconds. We repeat the experiment 10 times for statistical convergence.

C. Discussion of Results

We show the Y-PSNR vs. channel SNR for different cases in Figs. 7 to 10, where the solid lines correspond to simulation results and dashed lines correspond to theoretical results. In Fig. 7, the video is *skate* and the energy constraint for a GOP is $\mathcal{E}_c = 20$ mJ. The maximum energy that a GOP can possibly use with BPSK, FEC rate r_{min}, maximum power S_{max} and N_{max} transmissions is 75 mJ. For all cases, the Y-PSNR increases with channel SNR. When using the proposed UEP, the gap between

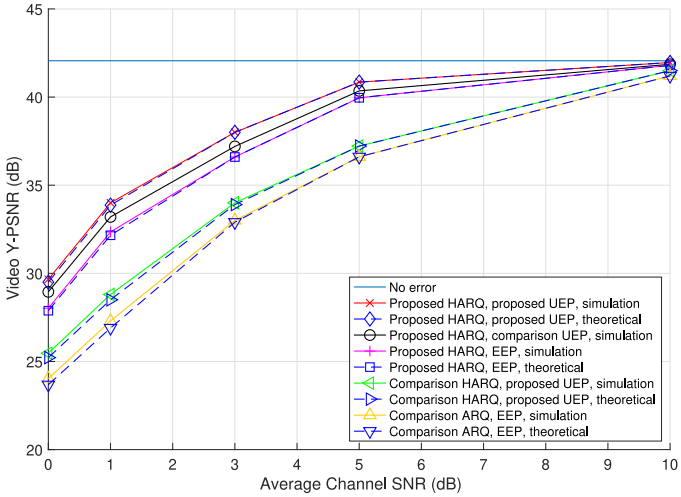


Fig. 7. Y-PSNR for *skate*. High energy constraint.

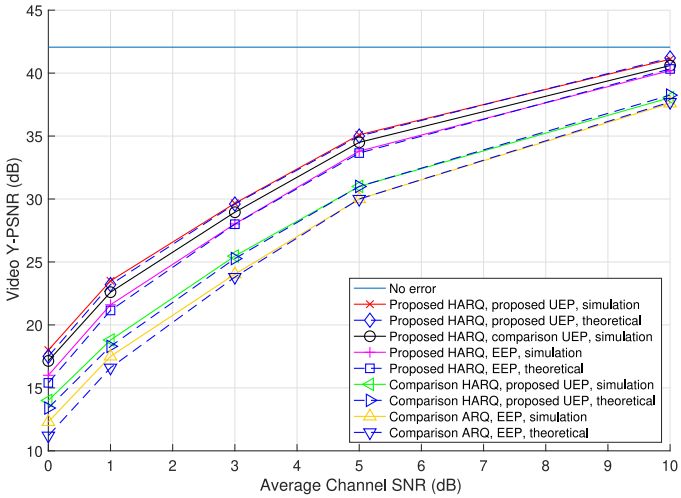


Fig. 8. Y-PSNR for *skate*. Low energy constraint.

the proposed HARQ and comparison HARQ is about 4 dB in the low SNR region, which shows a significant advantage for the proposed HARQ. When using the proposed HARQ, the gap between the proposed UEP and EEP is about 1.6 dB in the low SNR region, and the gap between the proposed UEP and comparison UEP is about 0.7 dB in the low SNR region. The 0.7 dB gap means that it is not desirable to divide the whole GOP into three groups as in the comparison UEP. But some conventional UEP schemes, e.g., those which use different FEC for different parts of the video data, only allow a few protection levels due to the complexity. However, the proposed UEP is able to assign different protection for each frame, thus achieving higher video quality. In Fig. 8, the energy constraint $\mathcal{E}_c = 10$ mJ. The curves show a similar trend to Fig. 7. In both figures, the theoretical analysis and simulation are close except for the low SNR region. This is because the PER is large in the low SNR region, so the model is less accurate because the errors that occur in a frame, but which propagate from different frames, are not uncorrelated.

Figs. 9 and 10 show the Y-PSNR for *football* and *interview* with energy constraint $\mathcal{E}_c = 20$ mJ. Comparing Figs. 9 and

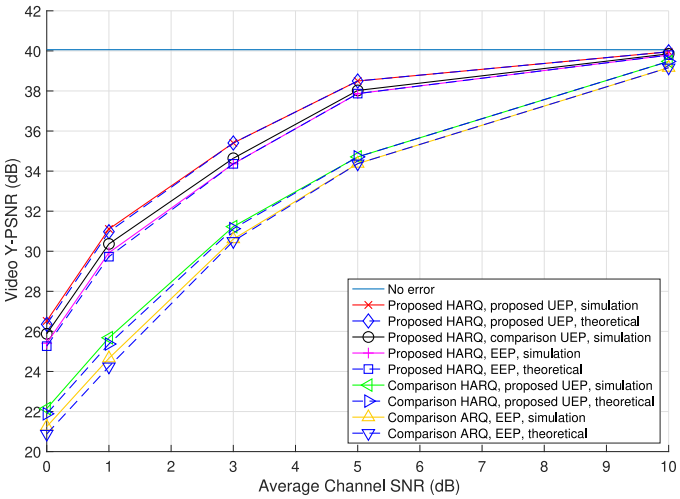


Fig. 9. Y-PSNR for *football*. High energy constraint.

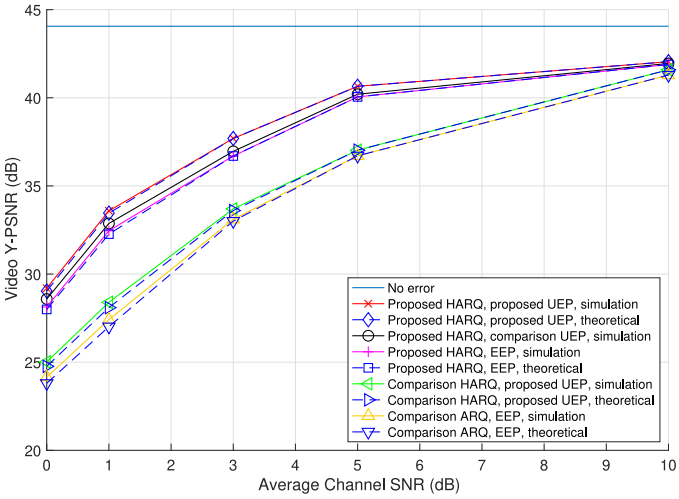


Fig. 10. Y-PSNR for *interview*. High energy constraint.

10 to 7, we find that for a given scheme, Y-PSNR depends on the video motion. With higher motion, Y-PSNR is lower because frame copy is less effective for high motion video. The difference between UEP and EEP, which is defined as the UEP gain, depends on video motion. With higher motion, the gain is smaller. However, the difference between the proposed HARQ and comparison HARQ is almost the same for all videos.

Next we evaluate the influence of the parameters σ_0^2 and α, which depend on the video, on the UEP gain. Let the PER sequence with respect to EEP be $P_{e0}, P_{e0}, ..., P_{e0}$, and let the PER sequence with respect to UEP be $P_{e1}, P_{e2}, ..., P_{eT}$. Then the UEP gain (dB) is

$$10log_{10}\left(\frac{\sigma_0^2\sum_{i=1}^{T}P_{e0}\sum_{\tau=0}^{T-i}\frac{1}{1+\alpha\tau}}{\sigma_0^2\sum_{i=1}^{T}P_{ei}\sum_{\tau=0}^{T-i}\frac{1}{1+\alpha\tau}}\right)$$
$$=10log_{10}\left(\frac{P_{e0}\sum_{i=1}^{T}\sum_{\tau=0}^{T-i}\frac{1}{1+\alpha\tau}}{\sum_{i=1}^{T}P_{ei}\sum_{\tau=0}^{T-i}\frac{1}{1+\alpha\tau}}\right). \quad (18)$$

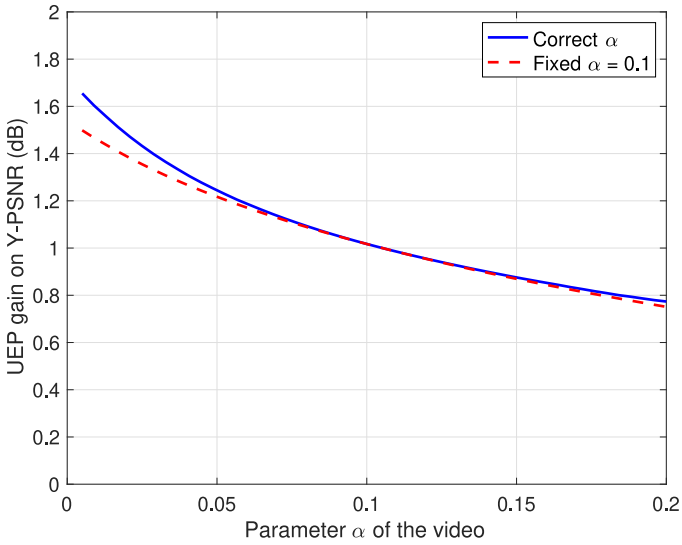


Fig. 11. UEP gain on Y-PSNR vs. α.

From Equations (16) and (17), we find that the unequal PER sequence $P_{e1}, P_{e2}, ..., P_{eT}$ does not depend on σ_0^2, since σ_0^2 is absorbed by the Lagrangian multiplier λ. Thus, the UEP gain in Equation (18) only depends on α. The α values for *football*, *skate* and *interview* are 0.18, 0.0065, 0.005, respectively. In Fig. 11, the blue solid curve shows the UEP gain vs. video parameter α. We curve fit the parameter α based on the video and use this correct value for the problem solving. The gain decreases with the increase of α, and high motion video has large α, so the gain is smaller for high motion video. This is also demonstrated from Figs. 7 to 10. The red dashed curve shows the UEP gain vs. video parameter α, but fixed $\alpha = 0.1$ is used for problem solving. There is a small loss due to unmatched parameter α for low motion video, and the loss is negligible for medium and high motion video.

Comparing Figs. 7, 9 and 10, we find that the performance gain for the proposed HARQ over the comparison HARQ is similar for videos with different degrees of motion. This can be demonstrated by Equation (4). Supposing EEP is used, we let the PER sequence with respect to the comparison HARQ be $P_{e0}, P_{e0}, ..., P_{e0}$, and let the PER sequence with respect to the proposed HARQ be $P'_{e0}, P'_{e0}, ..., P'_{e0}$. Then the HARQ gain (dB) when using EEP is

$$10log_{10}\left(\frac{\sigma_0^2\sum_{i=1}^{T}P_{e0}\sum_{\tau=0}^{T-i}\frac{1}{1+\alpha\tau}}{\sigma_0^2\sum_{i=1}^{T}P'_{e0}\sum_{\tau=0}^{T-i}\frac{1}{1+\alpha\tau}}\right)$$
$$=10log_{10}\left(\frac{P_{e0}}{P'_{e0}}\right) \quad (19)$$

This gain does not depend on σ_0^2 and α. When using UEP, this gain is also largely independent of σ_0^2 and α according to numerical results. Thus, the gain of the proposed HARQ system is almost constant for different videos, and fixed parameters provide close-to-optimal performance. This means the proposed algorithm has the important advantage that if optimization results are obtained offline, then minimal computation is caused

by the proposed algorithm during transmission. This reduces the computational power and delay, and allows the algorithm to be real-time.

VII. Conclusion

In this paper, we investigate energy-optimized wireless video transmission employing Hybrid ARQ and AMC. We consider both the unequal importance for multiple transmissions of the same video content, and the unequal importance of different video contents. We divide the problem into two sub-problems which solve the unequal importance for multiple transmissions and solve the UEP for different video content, respectively. The unequal importance for multiple transmissions shows that the transmissions should be sent with increasing order of cost to minimize the overall energy consumption. The video UEP shows that the frames in a GOP should have increasing order of packet error rate to minimize the MSE in the GOP. We compare the proposed scheme to the HARQ that uses single alphabet size and variable power, whereas the proposed scheme uses variable alphabet size and variable power. Compared to the video UEP scheme which is only allowed to have three (or few) levels of protections, the proposed scheme is able to assign optimal protection for each frame in the GOP. Simulations show that in the low SNR region, the proposed scheme outperforms the comparison HARQ by about 4 dB, and outperforms the comparison video UEP scheme by 0.8 to 1.6 dB depending on the video motion. There is a total gain of 4.8 to 5.6 dB compared to video transmission using conventional HARQ without any video UEP.

References

[1] "Global mobile data traffic forecast update." 2019. [Online]. Available: http://goo.gl/ylTuVx

[2] R. Trestian, A.-N. Moldovan, O. Ormond, and G.-M. Muntean, "Energy consumption analysis of video streaming to android mobile devices," in *Proc. IEEE Netw. Operations Manage. Symp.*, 2012, pp. 444–452.

[3] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.

[4] W. Su, S. Lee, D. A. Pados, and J. D. Matyjas, "Optimal power assignment for minimizing the average total transmission power in hybrid-ARQ Rayleigh fading links," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1867–1877, Jul. 2011.

[5] T. V. Chaitanya and E. G. Larsson, "Optimal power allocation for hybrid ARQ with chase combining in IID Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1835–1846, May 2013.

[6] H. Seo and B. G. Lee, "Optimal transmission power for single-and multi-hop links in wireless packet networks with ARQ capability," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 996–1006, May 2007.

[7] C. Shen, T. Liu, and M. P. Fitz, "On the average rate performance of hybrid-ARQ in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3339–3352, Nov. 2009.

[8] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, "Delay constrained rate and power adaptation over correlated fading channels," in *Proc. IEEE Global Telecommun. Conf.*, 2004, pp. 3448–3453.

[9] B. Makki and T. Eriksson, "On hybrid ARQ and quantized CSI feedback schemes in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 986–997, Apr. 2012.

[10] H. Jin, C. Cho, N.-O. Song, and D. K. Sung, "Optimal rate selection for persistent scheduling with HARQ in time-correlated Nakagami-m fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 637–647, Feb. 2011.

[11] P. Larsson, L. K. Rasmussen, and M. Skoglund, "Throughput analysis of hybrid-ARQ–A matrix exponential distribution approach," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 416–428, Jan. 2016.

[12] T. Villa, R. Merz, R. Knopp, and U. Takyar, "Adaptive modulation and coding with hybrid-ARQ for latency-constrained networks," in *Proc. 18th Eur. Wireless Conf. (Eur. Wireless)*, 2012, pp. 1–8.

[13] C. Ji, D. Wang, N. Liu, and X. You, "On power allocation for incremental redundancy Hybrid ARQ," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1506–1518, Mar. 2015.

[14] M. Jabi, L. Szczecinski, M. Benjillali, and F. Labeau, "Outage minimization via power adaptation and allocation in truncated hybrid ARQ," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 711–723, Mar. 2015.

[15] X. Zhang and Q. Du, "Adaptive low-complexity erasure-correcting code-based protocols for QoS-driven mobile multicast services over wireless networks," *IEEE Trans. Veh. Technol.*, vol. 55, no. 5, pp. 1633–1647, Sep. 2006.

[16] F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Rate-distortion optimized hybrid error control for real-time packetized video transmission," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 40–53, Jan. 2006.

[17] A. Aqil, A. O. Atya, S. V. Krishnamurthy, and G. Papageorgiou, "Streaming lower quality video over LTE: How much energy can you save?" in *Proc. IEEE 23rd Int. Conf. Netw. Protocols*, 2015, pp. 156–167.

[18] M. E. El Dien, A. A. Youssif, and A. Z. Ghalwash, "Energy efficient and QoS aware framework for video transmission over wireless sensor networks," *Wireless Sensor Netw.*, vol. 8, no. 03, pp. 25–36, 2016.

[19] T. Maksymyuk, L. Han, X. Ge, H.-H. Chen, and M. Jo, "Quasi-quadrature modulation method for power-efficient video transmission over LTE networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2083–2092, Jan. 2014.

[20] J. Wu, B. Cheng, M. Wang, and J. Chen, "Energy-efficient bandwidth aggregation for delay-constrained video over heterogeneous wireless networks." *IEEE J. Select. Areas Commun.*, vol. 35, no. 1, pp. 30–49, Jan. 2017.

[21] Q. Wang, T. M. Steinman, and W. Wang, "Quality driven modulation rate optimization for energy efficient wireless video relays," *Comput. Commun.*, vol. 115, pp. 2–9, 2018.

[22] C. Singhal, S. De, R. Trestian, and G.-M. Muntean, "Joint optimization of user-experience and energy-efficiency in wireless multimedia broadcast," *IEEE Trans. Mobile Comput.*, vol. 13, no. 7, pp. 1522–1535, Jul. 2014.

[23] T. Q. Duong, N.-S. Vo, T.-H. Nguyen, M. Guizani, and L. Shu, "Energy-aware rate and description allocation optimized video streaming for mobile D2D communications," in *Proc. IEEE Int. Conf. Commun.*, 2015, pp. 6791–6796.

[24] L. Liu, Y. Yi, J.-F. Chamberland, and J. Zhang, "Energy-efficient power allocation for delay-sensitive multimedia traffic over wireless systems," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2038–2047, Jan. 2014.

[25] Q. Wang, W. Wang, S. Jin, and H. Zhu, "Cross-layer source-channel control for future wireless multimedia services: Energy, latency, and quality investigation," *IET Commun.*, vol. 11, no. 17, pp. 2575–2584, Nov. 2017.

[26] X. Ge *et al.*, "Energy-efficiency optimization for MIMO-OFDM mobile multimedia communication systems with QoS constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2127–2138, Jun. 2014.

[27] Z. Ye, R. Hegazy, W. Zhou, P. Cosman, and L. Milstein, "Joint energy optimization of video encoding and transmission," in *Proc. Picture Coding Symp.*, 2018, pp. 116–120.

[28] A. A. Khalek, C. Caramanis, and R. Heath, "A cross-layer design for perceptual optimization of H.264/SVC with unequal error protection," *IEEE J. Select. Areas Commun.*, vol. 30, no. 7, pp. 1157–1171, Aug. 2012.

[29] R. Xiong, D. S. Taubman, and V. Sivaraman, "PET protection optimization for streaming scalable videos with multiple transmissions," *IEEE Trans. Image Process.*, vol. 22, no. 11, pp. 4364–4379, Nov. 2013.

[30] M. van der Schaar and D. S. Turaga, "Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission," *IEEE Trans. Multimedia*, vol. 9, no. 1, pp. 185–197, Jan. 2007.

[31] H.-C. Wei, Y.-C. Tsai, and C.-W. Lin, "Prioritized retransmission for error protection of video streaming over WLANs," in *Proc. Int. Symp. Circuits Syst.*, 2004, vol. 2, pp. II–65.

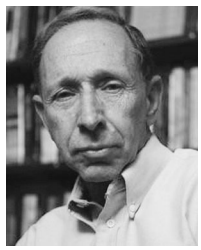
[32] T. P. Fowdur, D. Indoonundon, and S. K. Soyjaudah, "An enhanced framework for H.264 video transmission with joint prioritisation of retransmission and concealment order," in *Proc. 9th Int. Symp. Commun. Syst., Netw. Digit. Signal Process.*, 2014, pp. 634–639.

[33] V. Vadori, A. V. Guglielmi, and L. Badia, "Markov analysis of video transmission based on differential encoded HARQ," in *Proc. IEEE 17th Int. Symp. World Wireless, Mobile Multimedia Netw.*, 2016, pp. 1–9.

[34] B. Tirouvengadam, R. Radhakrishnan, and A. Nayak, "CAAHR: Content aware adaptive HARQ retransmission scheme for 4G/LTE network," in *Proc. 4th Int. Conf. Ubiquitous Future Netw.*, 2012, pp. 456–461.
[35] A. Rapaport, W. Liu, L. Ma, G. S. Sternberg, A. Ziera, and A. Balasubramanian, "Adaptive HARQ and scheduling for video over LTE," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2013, pp. 1584–1588.
[36] C. Zhu, Y. Huo, B. Zhang, R. Zhang, M. El-Hajjar, and L. Hanzo, "Adaptive-truncated-HARQ-aided layered video streaming relying on interlayer FEC coding," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1506–1521, Mar. 2016.
[37] D. T. Nguyen and J. Ostermann, "Congestion control for scalable video streaming using the scalability extension of H.264/AVC," *IEEE J. Select. Topics Signal Process.*, vol. 1, no. 2, pp. 246–253, Aug. 2007.
[38] T. Gan, L. Gan, and K.-K. Ma, "Reducing video-quality fluctuations for streaming scalable video using unequal error protection, retransmission, and interleaving," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 819–832, Apr. 2006.
[39] L. Han, S.-S. Kang, H. Kim, and H. P. In, "Adaptive retransmission scheme for video streaming over content-centric wireless networks," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1292–1295, Jun. 2013.
[40] H. Mukhtar, A. Al-Dweik, M. Al-Mualla, and A. Shami, "Low complexity power optimization algorithm for multimedia transmission over wireless networks," *IEEE J. Select. Topics Signal Process.*, vol. 9, no. 1, pp. 113–124, Feb. 2015.
[41] R. Radhakrishnan and A. Nayak, "Cross layer design for efficient video streaming over LTE using scalable video coding," in *Proc. IEEE Int. Conf. Commun.*, 2012, pp. 6509–6513.
[42] J. Yang, Q. Zheng, H. Xi, and L. Hanzo, "Receiver-driven adaptive enhancement layer switching algorithm for scalable video transmission over link-adaptive networks," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 47–50, Jan. 2013.
[43] H. R. Ghaeini, B. Akbari, B. Barekatain, and A. Trivino-Cabrera, "Adaptive video protection in large scale peer-to-peer video streaming over mobile wireless mesh networks," *Int. J. Commun. Syst.*, vol. 29, no. 18, pp. 2580–2603, 2016.
[44] J.-H. Lee, G.-S. Hong, Y.-W. Lee, C.-K. Kim, N. Park, and B.-G. Kim, "Design of efficient key video frame protection scheme for multimedia Internet of Things (IoT) in converged 5G network," *Mobile Netw. Appl.*, vol. 24, pp. 208–220, Feb. 2019.
[45] W. Wang and A. Kwasinski, "Adaptive learning for scalable video transmission with HARQ over dynamic wireless channels," in *Proc. IEEE Int. Conf. Commun.*, 2015, pp. 3094–3099.
[46] F. Marx and J. Farah, "A novel approach to achieve unequal error protection for video transmission over 3G wireless networks," *Signal Process. Image Commun.*, vol. 19, no. 4, pp. 313–323, 2004.
[47] ITU-T and ISO/IEC JTC 1, "Advanced video coding for generic audiovisual services," *ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC)*, Mar. 2009.
[48] B. Girod and N. Fiarber, "Compressed video over networks," Telecommun. Lab., Univ. Erlangen-Nuremberg, Erlangen, Germany, 1999.
[49] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jan. 2000.
[50] L. Szczecinski, S. R. Khosravirad, P. Duhamel, and M. Rahman, "Rate allocation and adaptation for incremental redundancy truncated HARQ," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2580–2590, Jun. 2013.
[51] M. Jabi, M. Benjillali, L. Szczecinski, and F. Labeau, "Energy efficiency of adaptive HARQ," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 818–831, Feb. 2016.
[52] B. Zhang, P. Cosman, and L. Milstein, "Energy optimization for HARQ and AMC," in *Proc. 51st Asilomar Conf. Signals, Syst. Comput.*, 2017, pp. 1796–1800.
[53] D. Ghadiyaram, J. Pan, and A. C. Bovik, "A subjective and objective study of stalling events in mobile streaming videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 1, pp. 183–197, Jan. 2019.
[54] J. P. D. Ghadiyaram and A. C. Bovik, "LIVE mobile stall video database-II." 2016. [Online]. Available: http://live.ece.utexas.edu/research/LIVEStallStudy/index.html

Bentao Zhang received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2014, and the M.S. degree in electrical engineering in 2018 from the University of California at San Diego, La Jolla, CA, USA, where he is currently working toward the Ph.D. degree. His research interests include wireless communications and wireless video transmission.

Pamela Cosman (S'88–M'93–SM'00–F'08) received the B.S. degree (with Hons.) in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 1987, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1993. Following an NSF Postdoctoral fellowship with Stanford and with the University of Minnesota (1993–1995), she joined the faculty of the Department of Electrical and Computer Engineering, the University of California, San Diego, where she is currently working as a Professor. She has authored and coauthored more than 250 technical papers as well as one children's book, *The Secret Code Menace*, that introduces error correction coding through a fictional story. Her research interests include image and video compression and processing, and wireless communications. She received the ECE Departmental Graduate Teaching Award, Career Award from the National Science Foundation, Globecom 2008 Best Paper Award, HISB 2012 Best Poster Award, 2016 UC San Diego Affirmative Action and Diversity Award, and 2017 Athena Pinnacle Award (Individual in Education). Her administrative positions include serving as the Director of the Center for Wireless Communications (2006–2008), ECE Department the Vice-Chair (2011–2014), and an Associate Dean for Students (2013–2016). She has been a member of the Technical Program Committee or the Organizing Committee for numerous conferences, including most recently serving as Technical Program Co-Chair of ICME 2018. She was an Associate Editor of the IEEE Communications Letters (1998–2001), and an Associate Editor of the IEEE Signal Processing Letters (2001–2005). She was the Editor-in-Chief (2006–2009) as well as a Senior Editor (2003–2005, 2010–2013) of the IEEE Journal on Selected Areas in Communications. She is a member of Tau Beta Pi and Sigma Xi.



Laurence B. Milstein (S'66–M'68–SM'77–F'85) received the B.E.E degree from the City College of New York, New York, NY, USA, in 1964, and the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, USA, in 1966 and 1968, respectively. From 1968 to 1974, he was with the Space and Communications Group of Hughes Aircraft Company, and from 1974 to 1976, he was a member of the Department of Electrical and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. Since 1976, he has been with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, where he is the Ericsson Professor of Wireless Communications and former Department Chairman, working in the area of digital communication theory with special emphasis on spread-spectrum communication systems. He has also been a Consultant to both government and industry in the areas of radar and communications. He was an Associate Editor for Communication Theory for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor for book reviews for the IEEE TRANSACTIONS ON INFORMATION THEORY, an Associate Technical Editor for the IEEE Communications Magazine, and the Editor-in-Chief of the IEEE Journal of Selected Areas in Communications. He was the Vice President for Technical Affairs in 1990 and 1991 of the IEEE Communications Society, and is a former Chair of the IEEE Fellow Selection Committee. He is a recipient of the 1998 Military Communications Conference Long Term Technical Achievement Award, an Academic Senate 1999 UCSD Distinguished Teaching Award, an IEEE Third Millennium Medal in 2000, the 2000 IEEE Communication Society Armstrong Technical Achievement Award, and the 2002 MILCOM Fred Ellersick Award.