

Layered Wireless Video Relying on Minimum-Distortion Inter-Layer FEC Coding

Yongkai Huo, Mohammed El-Hajjar, Robert G. Maunder, and Lajos Hanzo, *Fellow, IEEE*

Abstract—Layered video coding is capable of progressively refining the reconstructed video quality with the aid of multiple layers of unequal importance. When the base layer (BL) is corrupted or lost due to channel impairments, the enhancement layers (ELs) must be discarded by the video decoder, regardless whether they are perfectly decoded or not, which implies that the transmission power assigned to the ELs is wasted. To circumvent this problem, we proposed a bit-level inter-layer forward error correction (IL-FEC) scheme for layered video transmission in our previous work, which implanted the systematic bits of the BL into the systematic bits of the ELs using exclusive-OR operations (XOR). This allowed the receiver to exploit the implanted bits of the ELs for assisting the BL's decoding and hence improved the overall system performance of our IL-FEC aided layered video scheme. In this treatise, we find the specific FEC coding rates in a real-time on-line fashion for the sake optimizing the overall system performance. The proposed procedure is widely applicable to diverse wireless transceivers and FEC codecs. Our simulation results show that the proposed optimized IL-FEC system outperforms the traditional optimal UEP by about 1.9 dB of E_b/N_0 at a peak signal-to-noise ratio (PSNR) of 38 dB. Viewing the improvements in terms of the video quality, 3.3 dB of PSNR improvement is attained at an E_b/N_0 of 10 dB, when employing a recursive systematic convolutional (RSC) code.

Index Terms— [Author, please supply index terms/keywords for your paper. To download the IEEE Taxonomy go to http://www.ieee.org/documents/2009Taxonomy_v101.pdf.]

I. INTRODUCTION

IN RECENT years, low-delay lip-synchronized wireless video communications became a reality and the road to this is detailed in [1]. Layered video coding [2] has been widely adopted by the existing video coding standards [3]–[6], which is capable of generating multiple layers of unequal importance. The most important layers are referred to as the base layer (BL), while the less important layers depend on the BL and are

referred to as enhancement layers (ELs). As an inherent component of layered video coding standards, partitioned mode video coding [5] encodes each video-frame slice into at most three partitions, namely partition A, B and C according to the error sensitivity of the bistreams, where partition A may be deemed to be the BL, while partitions B and C are treated as ELs. Furthermore, the moving picture expert group (MPEG)'s [1] developed a multiview profile (MVP) [3], where the left view and right view are encoded as a BL and an EL, respectively. Another recent layered video coding standard, referred to as scalable video coding (SVC) [4], [5] was designed as an extension of H.264/AVC [5]. In SVC, a video sequence is encoded into a stream containing multiple dependent layers, where a number of subset layers may be extracted and hence the bitrate may be controlled according to the users' specific requirements in practical video services. A common characteristic of layered video coding scheme is that the ELs may be assigned a lower priority than the BL and hence may be discarded in the scenario of network congestion or buffer overflow [7]. Moreover, when the related BL is corrupted or lost due to channel impairments, the layered video decoder has to discard the corresponding ELs, regardless whether they are correctly recovered or not.

Unequal error protection (UEP) was first proposed in [8], where a stronger forward error correction (FEC) code was allocated to the more important data and a weaker FEC code was allocated to the less important data. In [9] UEP was achieved by optimizing the transmission power assigned to individual source bits according to their importance. Additionally, UEP using convolutional codes was investigated in [10], where the rate-compatible convolutional codes (RCPC) proposed in [11] were employed. Moreover, a number of UEP techniques based on low-density parity-check (LDPC) codes were studied in [12]–[15], where the more important bits were more strongly protected than the less important bits. Moreover, most of the above-mentioned UEP contributions considered artificial signals, rather than realistic video signals, which may not be directly extended to the scenario of flawless, low-delay, lip-synchronized video communications.

In layered video coded bitstreams different layers have a different importance, which implies that employing UEP techniques would substantially improve the reconstructed video quality. Generally, these UEP techniques [16]–[31] may be classified into two categories,¹ namely the packet-level schemes [18], [20], [23]–[27], [30] and bit-level schemes [16], [17], [19], [21], [22], [28], [29], [31]. Specifically, the packet-level contributions [18], [20], [23]–[27], [30] mitigate the packet loss

Manuscript received June 02, 2013; revised November 04, 2013; accepted November 30, 2013. This work was supported by the EU's Concerto project, of the EPSRC under the auspices of the India-UK Advanced Technology Centre (IU-ATC) and that of the ERC's Advanced Fellow Grant. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Wenwu Zhu.

The authors are with the School of Electronics and Computer Science, University of Southampton, U.K. (e-mail: yh3g09@ecs.soton.ac.uk; meh@ecs.soton.ac.uk; rm@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2014.2300449

¹Thanks to an anonymous reviewer for this valuable suggestion.

events as exemplified by the packets lost in Internet-routers [32], while the bit-level ones are devoted to eliminating bit-errors of wireless scenarios [33]. Let us continue by considering the family of packet-level UEP schemes first [18], [20], [25], [26]. An UEP scheme was conceived for object-based video communications in [18] for achieving the best attainable system performance under specific bitrate and delay constraints in an error-prone network environment. A novel UEP method using Reed-Solomon codes was proposed in [20] for SVC video transmission over networks inflicting packet-loss events. Firstly, the layer-weighted expected zone of error propagation (LW-EZEP) was defined as an efficient performance metric for quantifying the error propagation effects imposed by packet loss events. Then, the corresponding Reed-Solomon (RS) coding rates were assigned based on LW-EZEP for minimizing the expected video distortion. Cross-layer operation aided scalable video streaming designed for error-prone channels was investigated in [25], where the RS coded UEP was optimized for robust video delivery. The expected video quality was evaluated based on both the available bandwidth and the packet loss ratio (PLR) encountered, which was then further improved by employing content-aware bitrate allocation. Finally, an error concealment method was invoked at the receiver. The authors of [26] studied an UEP scheme using Luby Transform (LT) codes for recovering the video packets lost owing to network congestions. The above-mentioned UEP methodologies tended to employ hard decoded FEC [34] codes and dispensed with soft-information exchange in the decoding process, since they are carried out at the application layer. Similarly, bit-level UEP techniques have also been widely investigated [16], [17], [19], [21], [22], [29]. In [17], an UEP scheme using a turbo transceiver was optimized for wireless video telephony. The UEP performance of data-partitioned [5] H.264/AVC video streaming systems using recursive systematic convolutional (RSC) codes was evaluated in [29], while turbo coded modulation based UEP was investigated in [19], where both the cutoff rates and the channel capacity of each of the UEP levels was determined. The authors of [21] considered the unequal importance of both the video-frames in a group of pictures (GOP) and the significance of the diverse macroblocks (MB) in a video frame for transmission over wireless channels, where a prompt and efficient fast rate allocation scheme was also investigated. However, only three protection classes were discussed in [21], which limits the attainable system performance. The same authors also considered the different importance of the intra-coded (I) frame and of the predicted (P) frames within a GOP in [22], where the video bits of different importance were mapped to the different-integrity modulation constellation points of hierarchical quadrature amplitude modulation (QAM). These bit-level schemes tended to employ physical layer FEC codes [35] and perform soft decoding [34] for wireless video communications.

In the traditional UEP scheme conceived for layered video communication, variable-rate FEC codes were assigned to the different-sensitivity layers for improving the reconstructed video quality. However, when the BL is corrupted or lost, the ELs have to be discarded by the video decoder, regardless whether they are perfectly decoded or not, which implies that

both the transmission power and the bandwidth assigned to the ELs is wasted. Hence it is beneficial improve the protection of the more important BL with the aid of the ELs. Hence, the authors of [24] proposed a packet-level solution based on fountain codes [23] for scalable video multicast over networks inflicting packet loss events, where ELs conveyed parity information protecting the more important BL. By contrast, hierarchical network coding [27] and generalized network coding [30] techniques were also proposed for combating packet-loss events, when transmitting information of unequal importance. Furthermore, a bit-level layer-aware FEC (LA-FEC) philosophy using a hard-decoded Raptor code was designed for scalable video transmission over the binary erasure channels (BEC) in [28], [36]. The Raptor encoder generated the parity bits right across the BL and the ELs at the transmitter. As a benefit, the parity bits of the ELs may be utilized for assisting in correcting the errors residing in the BL at the receiver. Motivated by these advances, we developed a bit-level inter-layer coded FEC (IL-FEC) scheme layered video telephony over wireless fading channels relying on a soft-decoded systematic FEC code in [31], where the systematic bits of the BL are implanted into the ELs at the transmitter. At the receiver, the BL's bits implanted into the ELs may be utilized for correcting the BL. The above-mentioned IL-FEC technique of [31] was also combined with the UEP philosophy for the sake of further improving the attainable system performance. In [31], a number of coding rates were tested for the sake of proving the benefits of the proposed IL technique, where the code rates arrangements were determined empirically. However, in practical scenarios, different configurations of video codecs and different video sequences may have different characteristics, which may require different channel coding rates for achieving the best system performance.

Against this background, in this treatise, we propose a technique for finding the optimized coding rates for coded bitstreams "on-the-fly" at the transmitter, which optimizes the IL-FEC coded system performance. Specifically, we find the coding rates achieving the minimum video quality distortion with the aid of the mutual information (MI) between the log-likelihood ratios (LLRs) and the corresponding video bits. In this context, the soft-decoding metric of the FEC codec and of the demodulator are characterized by lookup tables (LUTs), since these cannot be characterized theoretically. Finally, the widely employed SVC-H.264 codec will be utilized for the sake of benchmarking our proposed system. In a nutshell, we focus our attention on the optimization of bit-level IL-FEC encoded scalable video communications over wireless fading channels. Specifically, our proposed technique is significantly different from the existing contributions [16], [20], [21], [23], [25], [26], [28] in the following aspects. Firstly, our technique is proposed for layered video communication over wireless channels, while the packet-level contributions of [20], [23], [25], [26] consider networks, which inflict packet-loss events or on networks modeled by the BEC. Secondly, IL-FEC relies on bit-level soft-decoding aided channel decoders, such as an RSC code, while [20], [23], [25], [26], [28] considered hard-decoded Reed-Solomon or fountain codes. Moreover, we propose a solution employing near-continuous coding rates, while only

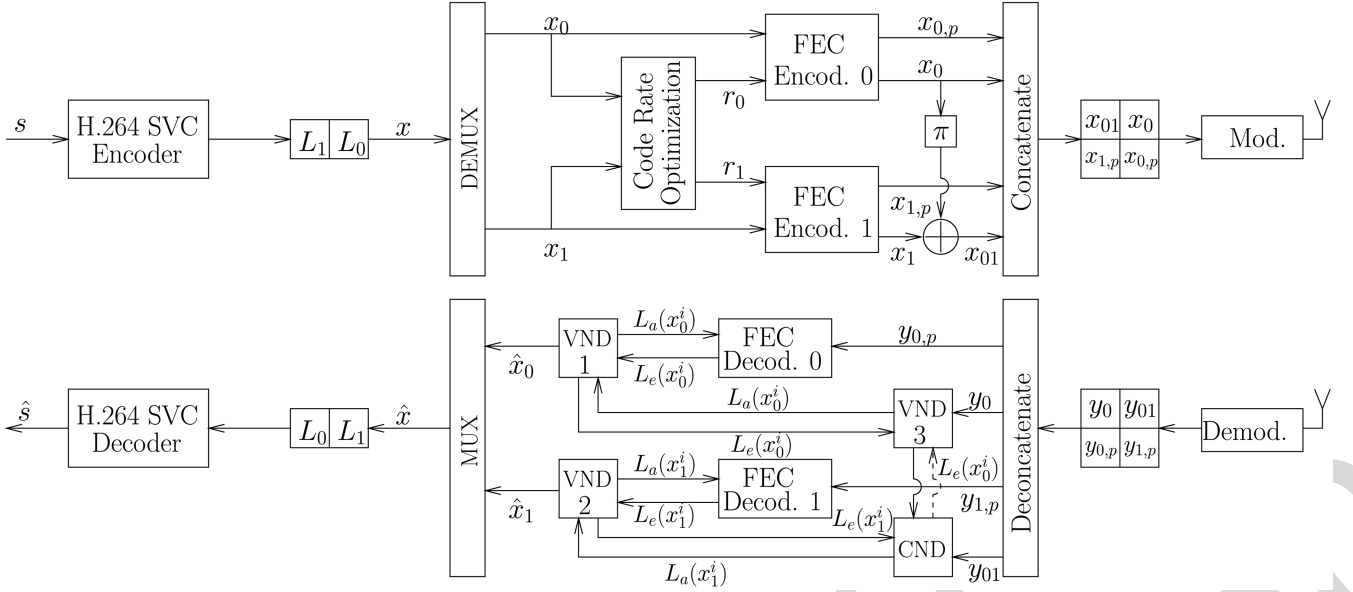


Fig. 1. Block diagram of the proposed IL-FEC coded SVC H.264/AVC coded video system, where r_0 and r_1 represent the code rates for FEC encoder 0 and 1, respectively. The “Code Rate Optimization” block will be detailed in Section III.

three protection rates - namely low-order, medium-order and high-order protection - were utilized in [16], [21], which limits the system’s performance. Finally, we find the optimized coding rates for the IL-FEC aided UEP system of [31], where the basic principle of IL-FEC was conceived, but the assisted focus and the system optimization was set aside for further research. Against this background, the rationale and novelty of this paper is summarized as follows.

- 1) *We set out to optimize soft-decoding bit-level IL-FEC encoded scalable video communication over wireless fading channels relying on a new objective function (OF) for the sake of finding the best code rate. Specifically, our technique relies on the MI and LUTs for finding the optimized coding rates in a real-time, online fashion, while minimizing video quality distortion. In contrast to the packet-level contributions of [20], [23], [25], [26], where the authors considered hard-decoded video streams transmitted over networks, which inflict packet-loss events or networks modelled by BEC, we consider transmission over non-dispersive Rayleigh fading wireless channels, where the layered video stream is protected by FEC codes and then soft-decoded at the receiver.*
- 2) *We utilize the MI and LUTs for the sake of optimizing the proposed system’s performance, where we show an E_b/N_0 gain of 1.9 dB at a PSNR of 38 dB. Alternatively, viewing the improvements in terms of the attainable video quality, or a PSNR gain of 3.3 dB is achieved at an E_b/N_0 of 10 dB over the benchmarker of [21].*
- 3) *A general solution is proposed for optimizing the coding rate allocation for arbitrary non-iterative FEC codes protecting, arbitrary layered video codecs. More explicitly, our solution is applicable to arbitrary transceivers provided that the receiver is capable of generating near-Gaussian distributed soft-information.*

We use the SVC-H.264/AVC scheme in our simulations, but our proposed scheme is not limited to SVC, it can be readily applied in any arbitrary system relying on layered video coding, such as the partitioned video coding regime of [31]. The rest of this paper is organized as follows. Section II details our proposed IL-FEC system model and the related video transmission techniques. Then the coding-rate optimization of the system is detailed in Section III. The performance of our optimized IL-FEC scheme using a RSC codec is benchmarked in Section IV using two video sequences having different motion characteristics. Finally, we offer our conclusions in Section V.

II. SYSTEM OVERVIEW

In this section, we briefly introduce the architecture of the inter-layer FEC scheme conceived for scalable video transmission, while further details on this technique may be found in [31]. The architecture of the system is displayed in Fig. 1, where the SVC mode of H.264 [5] is employed. Moreover, two layers, namely the BL L_0 and the EL L_1 are assumed to be generated by the SVC encoder, noting that more complex scenarios may be readily applied - we simply use two layers here for ease of explanation.² The variable node decoder (VND) and check node decoder (CND) [37] used in Fig. 1 are illustrated in Fig. 2. Both the VND and CND blocks accept a number of soft information inputs and generate the related soft information outputs for iteratively exploiting the deliberately imposed inter-dependencies between the coded layers L_0 and L_1 . Specifically, assuming that u_1 , u_2 and $u_3 = u_1 \oplus u_2$ are random binary variables, the VND of Fig. 2 sums two LLR inputs for generating an improved LLR output, which may be formulated as $L_{o_3}(u_1) = L_{i_1}(u_1) + L_{i_2}(u_1)$. Given the confidence of the bits u_1 and u_2 , the boxplus operation of $L(u_3 = u_1 \oplus u_2) = L(u_1) \boxplus L(u_2)$

²In our results in Section IV, we show how our proposed techniques work for SVC encoder using three layers.

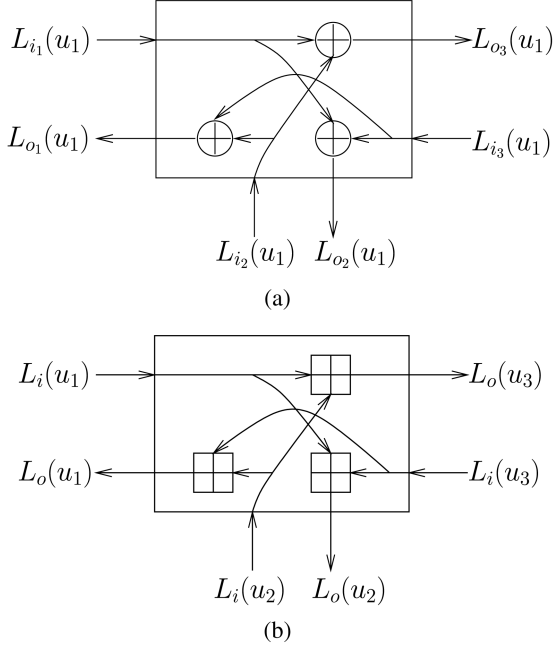


Fig. 2. The structure of VND and CND, where \oplus and \boxplus represent the addition and boxplus action, respectively. $L_i(\cdot)$ and $L_o(\cdot)$ represent the input and output LLR, respectively. (a) VND. (b) CND.

[38] may be utilized for deriving the confidence of the bit u_3 . Specifically, the boxplus operation \boxplus may be formulated as follows [39]

$$\begin{aligned} L(u_1) \boxplus L(u_2) &= \log \frac{1 + e^{L(u_1)} e^{L(u_2)}}{e^{L(u_1)} + e^{L(u_2)}} \\ &= \text{sign}[L(u_1)] \cdot \text{sign}[L(u_2)] \cdot \min[|L(u_1)|, |L(u_2)|] \\ &\quad + \log \left[1 + e^{-|L(u_1) + L(u_2)|} \right] \\ &\quad - \log \left[1 + e^{-|L(u_1) - L(u_2)|} \right]. \end{aligned} \quad (1)$$

In contrast to the VND block, the CND operation of Fig. 2 may be expressed as $L_o(u_3) = L_i(u_1) \boxplus L_i(u_2)$ for calculating the LLR of bit u_3 , given the LLR input of the bits u_1 and u_2 .

In Section II-A, the IL-FEC encoding techniques invoked at the transmitter are illustrated followed by our IL-FEC decoding techniques in Section II-B, where we emphasize on how the VND and the CND can exploit the IL information for improving the overall system performance. In Section II-A and II-B, we assume that the layers L_0 and L_1 contain an identical number of bits for simplifying our explanations, but our algorithm may be readily applied to the scenario of having an unequal number of bits as in [31].

A. Transmitter Model

At the transmitter of Fig. 1, the video source signal s is compressed using the SVC of the H.264 encoder, generating layers L_0 and L_1 . Then the output bitstream x is de-multiplexed into the bitstreams of x_0 and x_1 by the DEMUX block of Fig. 1, where x_0 and x_1 represent the bitstream of the BL L_0 and of the EL L_1 , respectively. Afterwards, the resultant two layers are encoded as follows:

- 1) The BL bit sequence x_0 representing L_0 will be encoded by the FEC encoder 0 of Fig. 1, which results in the encoded bits containing the systematic bits x_0 and parity bits $x_{0,p}$.
- 2) The bit sequence of the EL x_1 representing L_1 will firstly be encoded into the systematic bits x_1 and the parity bits $x_{1,p}$ by the FEC encoder 1. Then the XOR operation will be utilized for implanting the systematic information of x_0 into the systematic information of x_1 , which is carried out without changing the parity bits $x_{1,p}$ of the EL L_1 , as shown in Fig. 1. Specifically, the implantation process results in the check bits of $x_{01}^i = x_0^i \oplus x_1^i$. After this procedure, both the check bits x_{01}^i and the parity bits $x_{1,p}$ are output.

Finally, the bit sequences x_0 , $x_{0,p}$, x_{01} and $x_{1,p}$ are concatenated into a joint bitstream for transmission. Note however that the layers x_0 and x_1 may contain a different number of bits. Additionally, the interleavers π_1 are employed for interleaving the BL x_0 , before its XOR-based implantation into the EL x_1 . Following the IL-FEC encoding procedure, the resultant bits are modulated by the binary phase-shift keying (BPSK) modulator of Fig. 1 and then transmitted over *non-dispersive uncorrelated Rayleigh fading wireless channels*.

B. Receiver Model

In this section, we exemplify the IL decoding process using the BL L_0 and the EL L_1 . At the receiver,³ BPSK demodulation is performed, which generates the LLRs containing the systematic information y_0 , y_{01} and the parity information $y_{0,p}$ and $y_{1,p}$, for the layers L_0 and L_1 , respectively. Following the demodulator, the IL-FEC decoder of Fig. 1 is invoked for exchanging extrinsic information across the two layers. Firstly, the FEC decoder 0 will decode the received information y_0 and $y_{0,p}$ for estimating the LLRs of the bits x_0 of the BL L_0 . Then, the resultant extrinsic LLR information of the BL L_0 will be input to the “VND1-VND3-CND-VND2” block of Fig. 1 for extracting the *a-priori* LLRs $L_a(x_1^i)$ ⁴ of the EL L_1 , which is carried out by following the processing of the LLRs in the VND 1, VND 3, CND and VND 2 components of Fig. 2. Specifically, the “VND1-VND3-CND-VND2” block of Fig. 1 performs the following operations step-by-step:

- 1) VND 1 generates the information of BL L_0 for VND 3. The extrinsic LLR $L_e(x_0^i)$ generated by the FEC decoder 0 is input to the VND 1 block of Fig. 1, which extracts the extrinsic LLR information $L_e(x_0^i)$ and forwards it to the VND 3 block of Fig. 1. Since VND 1 has two input branches, it simply duplicates the soft information $L_e(x_0^i)$.
- 2) VND 3 generates the information of BL L_0 for CND. Generally, the summation of the soft information y_0 and the extrinsic LLR $L_e(x_0^i)$ from VND 1 will be output to CND.
- 3) CND generates the information of layer L_1 for VND 2. The LLR information $L_e(x_0^i)$ and the received check information y_{01} is input to the CND block of Fig. 1 for extracting the LLR information of the systematic bit x_1^i , namely the soft input $L_a(x_1^i)$ of VND 2.

³The deinterleavers π^{-1} is not shown in the receiver of Fig. 1 for the sake of simplifying the system architecture.

⁴The subscripts “a” and “e” in L_a and L_e stand for the *a priori* information and extrinsic information [40], respectively.

- 4) VND 2 generates the information of EL L_1 for FEC decoder 1. The LLR information $L_a(x_1^i)$ extracted by the CND is input to the VND 2 block of Fig. 1, which extracts the LLR information $L_a(x_1^i)$ input to the FEC decoder 1 of Fig. 1.

Then, the FEC decoder 1 of Fig. 1 will decode the EL L_1 with the aid of the resultant *a-priori* LLR $L_a(x_1^i)$ and of the soft parity information received from the channel, namely $y_{1,p}$ of Fig. 1. Similar to the “VND1-VND3-CND-VND2” process, the “VND2-CND-VND3-VND1” process is capable of generating the soft information for the BL. With the aid of the “VND1-VND3-CND-VND2” process and the “VND2-CND-VND3-VND1” process, the FEC decoder 0 and FEC decoder 1 may exchange extrinsic information with each other, which constitute an iteration, where the iterative IL decoding process continues until the affordable number of iterations is exhausted or the BL L_0 is perfectly recovered.

III. OPTIMIZED FEC CODING RATES

In this section, we focus our attention on the “Code Rate Optimization” block of our scalable video system seen in Fig. 1, where IL-FEC coded UEP is employed for the sake of achieving an improved video quality. Specifically, the “Code Rate Optimization” block of Fig. 1 characterizes the receiver of Fig. 1 and assists the transmitter in finding the specific FEC coding rates for the layers L_0 and L_1 by minimizing the video distortion. Additionally, although having only two layers is assumed in Section II, more complex coding scenarios may be readily considered, as evidenced by the results of Section IV. Moreover, in our system each layer will be mapped to a single packet. Specifically, in the SVC of H.264 coded video scenario, a network abstraction layer unit (NALU) is packed into a single packet, but the different NALUs may carry a different number of bits. Note that in practice several NALUs may be encapsulated into a single packet, depending on the specific requirements of the network.

Below, we will use a classic RSC codec for describing the proposed solution, which aims for minimizing the reconstructed video distortion at the receiver. This is achieved by characterizing the behavior of the receiver seen in Fig. 1. However, the employment of our proposed techniques is not limited to the RSC codec. Firstly, we define the following notations that will aid our analysis:

- $d(L_0)$: the video distortion induced by the corruption of the BL L_0 , which is measured using the peak signal-to-noise ratio (PSNR);
- $d(L_1)$: the PSNR distortion, when the BL L_0 is correct while the EL L_1 is corrupted;
- $|L_i|$: the length of the bitstream of layer L_i , $0 \leq i \leq 1$;
- R : the overall coding rate of the system shown in Fig. 1;
- r_i : the coding rate of layer L_i , $0 \leq i \leq 1$;
- $p(L_0)$: the packet error rate (PER) of layer L_0 ;
- $p(L_1)$: the PER of layer L_1 , when the layer L_0 is correctly decoded.

According to the IL-FEC decoding process detailed in Section II, both the FEC decoders 0 and 1 of Fig. 1 affect the PER $p(L_0)$, where FEC decoder 1 feeds back *a-priori*

information about BL L_0 to the FEC decoder 0 through the “VND2-CND-VND3-VND1” blocks of Fig. 1. Specifically, the performance of the FEC decoder 0 depends on SNR , r_0 , $|L_0|$, while the performance of the FEC decoder 1 depends on SNR , r_1 , $|L_1|$. Hence $p(L_0)$ depends on the parameters SNR , $|L_0|$, $|L_1|$, r_0 and r_1 , which can be expressed as

$$p(L_0) = f_0(SNR, |L_0|, |L_1|, r_0, r_1). \quad (2)$$

On the other hand, $p(L_1)$ purely depends on the FEC decoder 1, hence it is determined by the parameters SNR , $|L_1|$ and r_1 , which may be expressed as

$$p(L_1) = f_1(SNR, |L_1|, r_1). \quad (3)$$

Given the specific layers L_0 and L_1 , $d(L_i)$ may be calculated experimentally as the PSNR degrades due to the erasure of layer L_i [20]. Then the expected PSNR distortion induced by BL L_0 may be estimated as $p(L_0) \cdot d(L_0)$. Additionally, when BL L_0 is correctly decoded, the expected PSNR distortion induced by EL L_1 may be estimated as $p(L_1) \cdot d(L_1)$. Hence the expected distortion jointly induced by the decoded layers L_0 and L_1 , denoted as $E[d(L_0, L_1)]$, may be estimated as

$$E[d(L_0, L_1)] = p(L_0) \cdot d(L_0) + [1 - p(L_0)] \cdot p(L_1) \cdot d(L_1). \quad (4)$$

Based on the above discussions, the expected distortion $E[d(L_0, L_1)]$ may be calculated by following the dependency chains displayed in Fig. 3, which relies on the parameters SNR , $|L_0|$, $d(L_0)$, r_0 , $|L_1|$, $d(L_1)$ and r_1 . Among these parameters, SNR , $|L_0|$, $|L_1|$ are determined by the specific layers L_0 , L_1 and the specific transmit power, while $d(L_0)$ and $d(L_1)$ are gleaned from experiments [20]. Hence the coding rates r_0 , r_1 determine the distortion $E[d(L_0, L_1)]$ expected for the layers L_0 , L_1 . Therefore, the objective of the “Code Rate Optimization” block seen in Fig. 1 is to find the specific rates r_0 and r_1 capable of minimizing the expected distortion $E[d(L_0, L_1)]$ of Eq. (4), which may be expressed as

$$\arg \min_{r_0, r_1} \{E[d(L_0, L_1)]\}, \quad (5)$$

subject to the condition of

$$\frac{|L_0|}{r_0} + \frac{|L_1|}{r_1} = \frac{|L_0| + |L_1|}{R}. \quad (6)$$

As illustrated in Fig. 3, Eq. (4) relies on the PER definitions of Eqs. (2) and (3), which cannot be theoretically solved. Below we propose our solution of Eqs. (2) and (3) with the assistance of LUTs and the MI, where LUTs are employed for characterizing the system components of Fig. 1 and MI is employed for numerically quantifying the reliability of the LLRs. Specifically, Section III-A introduces the preliminaries of solving Eqs. (2) and (3), including the LUTs, the MI flow of the VND and CND, followed by the PER estimation of the BL and EL in Section III-B. Then, we exploit both the estimated PERs $p(\cdot)$ and the video distortions $d(\cdot)$ for determining the optimized coding rates in Section III-C. Finally, Section III-D discusses the overheads imposed by the proposed techniques.

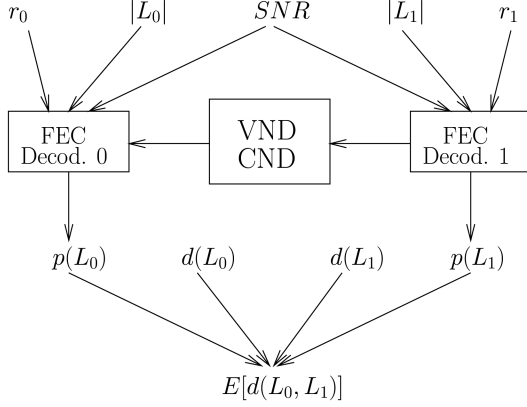


Fig. 3. Dependency chains for calculating the expected distortion formulated in Eq. (4).

A. Preliminaries

The receiver of Fig. 1 consists of the following components: demodulator, FEC decoder, VND, CND. The characteristics of these components jointly determine the PER $p(L_0)$ and $p(L_1)$, as shown in Fig. 3. However, neither the demodulator and nor the FEC decoder may be readily characterized theoretically for diverse system configurations, such as different transceivers, FEC generator polynomials, decoding metrics etc. In order to propose a more universal solution, below LUTs are employed for the sake of characterizing both the demodulator and the FEC decoder, while the MI flow of the VND and CND will be detailed afterwards.

1) *Lookup Tables*:: In our analysis we consider a RSC-coded BPSK-modulated system for transmissions over a *non-dispersive uncorrelated Rayleigh fading wireless channel*. Specifically, we model BPSK based transmissions over a Rayleigh fading wireless channel as a function of the channel SNR and generate the demodulator's output LLRs, where LLRs are quantified by the MI value, as in [41]. Furthermore, *since the MI invoked for quantifying the reliability of the soft information can be applied for Gaussian distributed LLRs [41] generated by arbitrary transceivers, our procedure may be deemed generically applicable, provided that the transceivers generate near-Gaussian distributed LLRs.*⁵ By contrast, the RSC codec is modeled as a function of the SNR, of the LLRs of the systematic bits and of the coding rate, while generating the RSC decoder's output extrinsic LLRs and estimating the PER of the RSC decoded packets, where the LLRs are quantified by the MI value. The following LUTs are created for modeling the transceiver and the RSC codec:

- $T_b(SNR)$: The MI value of the LLRs output by the BPSK demodulator for variable channel SNRs, as exemplified in Table IIa. Since this table relies on the SNR only, it may be stored in a one-dimensional space, where the LUTs' size requirements will be detailed in Section III-D.

⁵Our experiments, not included here for reasons of space-economy, suggest that higher order Quadrature Amplitude Modulation (QAM) can be readily applied for transmission over Rayleigh fading wireless channels, since it generates near-Gaussian distributed LLRs.

- $T_e(SNR, I_s, r)$: The MI value of the extrinsic LLR output of the RSC decoder recorded for variable channel SNRs, as exemplified in Table IIb, where I_s represents the MI value between the *a-priori* LLRs of the systematic information and the corresponding information bits, while r represents the coding rate of the RSC codec. Since this table relies on three parameters, it may be visualized in a three-dimensional space, where the LUTs' size requirements will also be detailed in Section III-D.
- $T_p(SNR, I_s, r)$: The PER value associated with the LLRs output by the RSC decoder at diverse channel SNRs, as exemplified in Table IIb, where I_s represents the MI between the *a-priori* LLRs of the systematic information and the corresponding information bits, while r represents the coding rate of the RSC codec. Note that this LUT may be generated simultaneously with the LUT $T_e(SNR, I_s, r)$, since they have the same input variables.

2) *MI Flow of the VND*:: As seen in Fig. 2a, the operation of the VND may be characterized as $L_{o3}(u_1) = L_{i1}(u_1) + L_{i2}(u_1)$. Assuming that the inputs $L_{i1}(u_1)$ and $L_{i2}(u_1)$ of the VND have the MI values of $I_{i1}[u_1; L_{i1}(u_1)]$ and $I_{i2}[u_1; L_{i2}(u_1)]$ respectively, the MI value of the output $L_{o3}(u_1)$ may be expressed as [42]

$$I_{o3}[u_1; L_{o3}(u_1)] = J \left(\sqrt{J^{-1}(I_{i1}[u_1; L_{i1}(u_1)])^2 + J^{-1}(I_{i2}[u_1; L_{i2}(u_1)])^2} \right), \quad (7)$$

where $J(\cdot)$ and $J^{-1}(\cdot)$ are defined in [37].

3) *MI Flow of the CND*:: Below, we firstly discuss the MI flow of the CND of Fig. 2b, which is designed for the layers L_0 and L_1 , both carrying the same number of bits. Then we will discuss the practical CND employed for realistic unequal length layers.

As seen in Fig. 2b, the action of the CND can be expressed as $L(u_3 = u_1 \oplus u_2) = L(u_1) \boxplus L(u_2)$ [38], where u_1, u_2 and $u_3 = u_1 \oplus u_2$ are random binary variables. Assuming that the inputs $L_i(u_1), L_i(u_2)$ of the CND have the MI values of $I_i[u_1; L(u_1)]$ and $I_i[u_2; L(u_2)]$ respectively, the MI value $I_o[u_3; L(u_3)]$ of u_3 has not been theoretically derived in the literature and since this is not the focus of our work, we model $I_o[u_3; L(u_3)]$ as

$$I_o[u_3; L(u_3)] = T_I(I_i[u_1; L(u_1)], I_i[u_2; L(u_2)]), \quad (8)$$

which is also characterized using LUT.

Given the MI $I_o[u_3; L(u_3)]$ in Eq. (8), we now discuss the practical scenario of the CND, where L_0 and L_1 may contain an unequal number of bits. As detailed in [31], the encoding architecture of L_0 and L_1 may be categorized into two scenarios, namely $|L_0| > |L_1|$ and $|L_1| > |L_0|$. Below we consider the scenario of generating $L(x_1^i)$, when $|L_0| > |L_1|$, which can be generalized to other scenarios. As detailed in [31], the bit x_{01}^i may be generated as

$$x_{01}^i = x_1^i \oplus \sum_{j=1}^n \oplus x_0^j, \quad (9)$$

where \oplus denotes the binary XOR operation. Additionally, since on average⁶ $|L_0|$ bits of L_0 were allocated to the $|L_1|$ bits of L_1 for generating the bits x_{01} , we have $n \in \{\lceil \frac{|L_0|}{|L_1|} \rceil, \lceil \frac{|L_0|}{|L_1|} \rceil - 1\}$ [31] for Eq. (9). More specifically, $(N - \frac{|L_0|}{|L_1|})$ proportion of the bits in the EL L_1 are encoded based on Eq. (9) using $n = N - 1$, while the remaining bits of the EL L_1 are encoded based on Eq. (9) relying on $n = N$, where we have $N = \lceil \frac{|L_0|}{|L_1|} \rceil$.

Based on Eq. (9), the corresponding LLRs $L(x_1)$ of the EL L_1 may be expressed as

$$L(x_1^i) = L(x_{01}^i) \boxplus \sum_{j=1}^n \boxplus L(x_0^j). \quad (10)$$

Then, the LLRs of $L(x_1)$ are constituted by a mixture of LLRs $L^N(x_1)$ and $L^{N-1}(x_1)$, which are calculated based on Eq. (10) in conjunction with $n = N$ and $n = N - 1$, respectively. We denote the MI values corresponding to the LLRs $L^N(x_1)$ and $L^{N-1}(x_1)$ as $I^N[x_1; L(x_1)]$ and $I^{N-1}[x_1; L(x_1)]$. Based on Eq. (8), the MI values $I^N[x_1; L(x_1)]$ and $I^{N-1}[x_1; L(x_1)]$ may be recursively calculated as in Eq. (11), where $1 \leq k \leq N$. Empirically, we assume that the LLRs $L^N(x_1)$ and $L^{N-1}(x_1)$ are near-Gaussian distributed, while their variance may be denoted as $J^{-1}(I^N[x_1; L(x_1)])$ and $J^{-1}(I^{N-1}[x_1; L(x_1)])$, respectively [37]. Again we assume that $L(x_1)$ exhibits near-Gaussian distribution, hence the variance of $L(x_1)$ may be formulated as shown in (11) and (12) at the bottom of the page.

Finally, the MI value of $L(x_1)$ may be expressed using the J function as in Eq. (13), shown at the bottom of the page.

⁶On average, each bit of L_1 is encoded by $\frac{|L_0|}{|L_1|}$ bits of L_0 for generating a single bit of x_{01} , as detailed in [31].

B. PER Estimation

The PER of the decoded BL and EL having coding rates of r_0 and r_1 will be estimated in Section III-B.1 and Section III-B.2, respectively.

1) *IL-FEC Coded BL*: As shown in Fig. 3, the PER of the BL L_0 depends on the characteristics of the IL-FEC decoder, including the generator polynomials, the trellis decoding techniques employed, as well as the VND and the CND of Fig. 2. Below we propose our solution conceived for estimating the PER of Eq. (3) with the aid of the LUTs defined in Section III-A.1.

The PER estimation flow-chart of the IL-FEC coded BL is illustrated in Fig. 4, which follows the decoding process detailed in Section II. Given the SNR , the MI value I_s of y_0 can be generated by the LUT $T_b(SNR)$. Furthermore, the MI value of the extrinsic LLR gleaned from the RSC decoder 0 may be expressed as $T_e[SNR, T_b(SNR), r_0]$. Then the “VND1-VND3-CND-VND2” process of Fig. 4 is capable of calculating the MI value $I_a(x_1)$ of $L_a(x_1^i)$. Afterwards, based on the LUT $T_e(SNR, I_s, r)$ of the RSC decoder 1, the MI value $I_e(x_1)$ of the extrinsic information may be readily expressed as

$$I_e(x_1) = T_e[SNR, T_b(SNR), r_1]. \quad (14)$$

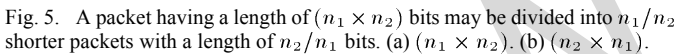
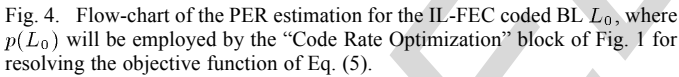
Then, following the “VND2-CND-VND3-VND1” process of Fig. 4, the improved MI value $I_a(x_0)$ of $L_a(x_0^i)$ can be calculated. Finally, the PER associated with the packet length of l ,⁷ denoted as $p(l)$, may be estimated as $T_p[SNR, I_a(x_0), r_0]$, which is also shown in Fig. 4. Below, we now detail the

⁷The packet containing l bits is employed for generating the LUT $T_p[SNR, I_a(x_0), r_0]$.

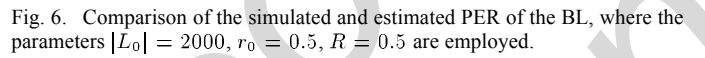
$$I^k[x_1; L(x_1)] = \begin{cases} T_I(I[x_0; L(x_0)], I[x_{01}; L(x_{01})]), & k = 1, \\ T_I(I^{k-1}[x_1; L(x_1)], T_I(I[x_0; L(x_0)], I[x_{01}; L(x_{01})])), & k > 1. \end{cases} \quad (11)$$

$$\begin{aligned} \sigma^2 &= D(L(x_1)) \\ &= \left(1 - N + \frac{|L_0|}{|L_1|}\right) \cdot D(L^N(x_1)) \\ &\quad + \left(N - \frac{|L_0|}{|L_1|}\right) \cdot D(L^{N-1}(x_1)) \\ &= \left(1 - N + \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^N[x_1; L(x_1)])^2 \\ &\quad + \left(N - \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^{N-1}[x_1; L(x_1)])^2. \end{aligned} \quad (12)$$

$$I[x_1; L(x_1)] = J(\sigma) = J\left(\sqrt{\left(1 - N + \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^N[x_1; L(x_1)])^2 + \left(N - \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^{N-1}[x_1; L(x_1)])^2}\right). \quad (13)$$



The burst error distribution of RSC codec has been investigated in [43], which is independent of the packet length. Let us consider a RSC-decoded packet having a length of $(n_1 \times n_2)$ bits. Then this packet may be partitioned in two ways, as displayed in Fig. 5. Specifically, it may be divided into n_1 packets,


$$\begin{aligned} p(L_1) &= f_1(SNR, |L_1|, r_1) \\ &= 1 - [1 - T_p(SNR, I_s, r_1)]^{|L_1|/l}, \end{aligned} \quad (18)$$

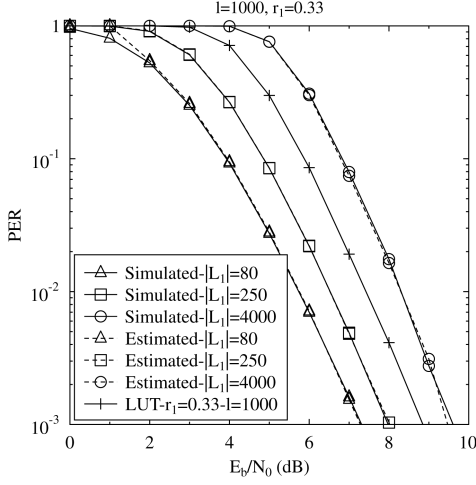


Fig. 7. Comparison of the simulated and estimated PER for the EL, where the parameters $l = 1000$, $r_1 = 0.33$ are employed.

where l is the packet length employed for generating the LUT T_p .

For visualizing the accuracy of our proposed EL PER estimation, we assume $r_1 = 0.33$, while the RSC codec is configured by the generator polynomials of [1011, 1101, 1111]. Moreover, we consider $|L_1|$ values ranging from 80 to 4000, while the LUT $T_p(SNR, I_s, r_1)$ is generated based on $l = 1000$. The resultant simulated and estimated PER of the EL L_1 is displayed in Fig. 7. Observe in Fig. 7 that the estimation error is much smaller than that in Fig. 6 due to the fact that the EL is protected by an RSC codec without using IL techniques.

C. Optimized Coding Rates

Based on Section III-B, the PERs of L_0 and L_1 are estimated according to the specified coding rates of r_0 and r_1 for the BL and EL, respectively. Hence the expected distortion $E[d(L_0, L_1)]$ imposed by the wireless transmission using rates of r_0 and r_1 may be readily calculated using Eq. (4). In order to find the rates r_0, r_1 minimizing the distortion $E[d(L_0, L_1)]$ in Eq. (5), a full search of the coding rate space may be performed. Note that r_1 is determined by Eq. (6) for a given r_0 , hence the full search is performed in a one-dimensional space to find the optimal r_0 .

To elaborate a little further, we consider the scenario of $SNR = 10$ dB, $R = 1/3$, $|L_0| = |L_1| = 100$, $d(L_0) = 20$, $d(L_1) = 10$. Based on the conditions formulated in Eq. (6), the coding rate r_1 of L_1 is determined as follows:

$$r_1 = \frac{|L_1|}{\frac{|L_0| + |L_1|}{R} - \frac{|L_0|}{r_0}} = \frac{r_0}{6 \cdot r_0 - 1} \quad (19)$$

$$r_0 \leq 1, r_1 \leq 1.$$

Then, based on Eqs. (2), (3) and (4) the expected value of $E[d(L_0, L_1)]$ in Eq. (5) may be formulated as

$$E[d(L_0, L_1)] = p(L_0) \cdot d(L_0) + [1 - p(L_0)] \cdot p(L_1) \cdot d(L_1)$$

$$= 20 \cdot f_0 \left(10, 100, 100, r_0, \frac{r_0}{6 \cdot r_0 - 1} \right)$$

$$+ 10 \cdot \left[1 - f_0 \left(10, 100, 100, r_0, \frac{r_0}{6 \cdot r_0 - 1} \right) \right]$$

$$\cdot f_1 \left(10, 100, \frac{r_0}{6 \cdot r_0 - 1} \right)$$

$$= g(r_0), \quad (20)$$

where the functions f_0 and f_1 are discussed in the ‘PER Estimation’ Section III-B. Note that we introduced the short-hand of $E[d(L_0, L_1)] = g(r_0)$ in Eq. (20) for simplicity. Based on Eq. (19), we can readily arrive at $\frac{1}{5} \leq r_0 \leq 1$. Then we artificially limit r_0 to the following discrete practical values $[0.2 : 0.05 : 1]$.⁸ Finally, the solution r_0 of the objective function $\arg \min \{E[d(L_0, L_1)]\}$ in Eq. (5) may be readily found to be:

$$\left\{ r_0 \mid \arg \min_{r_0} g(r_0), r_0 \in \{0.2 : 0.05 : 1\} \right\}, \quad (21)$$

which can be solved by performing a one-dimensional full search. Then, given r_0 the coding rate r_1 may be readily expressed from Eq. (19).

D. Overheads

All the optimization operations detailed in Section III are carried out at the transmitter. Below, we discuss the overheads imposed at the transmitter, while the overheads imposed at the receiver are given in [31]. The overheads imposed at the transmitter include the estimation of $d(\cdot)$, the generation of LUTs, the estimation of $p(\cdot)$, the imposed frame delay and the full search detailed in Section III-C. Among these overheads, the generation of LUTs only imposes extra off-line design-time, while the estimation of $d(\cdot)$, $p(\cdot)$ and the full search impose extra on-line run-time complexity. For notational simplicity, we use n_{snr}, n_{I_s}, n_r to denote the number of the parameters SNR, I_s, r , respectively, which are used for generating the LUTs. For example, if the values of I_s range from 0 to 1 with a step size of 0.01, we have $n_{I_s} = 101$. In the following, we will analyze these overheads in order to validate our system.

1) *Estimation of $d(\cdot)$* : $d(L_i)$ is estimated in a similar manner to the procedure of [20], where the distortion $d(L_0)$ may be obtained by decoding the bitstream in the presence of a corrupted BL L_0 . Alternatively, the solutions of [21], [25], [44] may be applied in our system.

2) *Generation of LUTs*: Three LUTs were generated in our proposed solution. The LUT $T_b(SNR)$ characterizes the channel and transceiver, hence this table has to be regenerated when the channel or the transceiver are changed. However, this LUT only has a single variable, hence it is straight-forward to generate the LUT $T_b(SNR)$. Additionally, the LUTs $T_e(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$ can be simultaneously generated by simulations, since they have the same variables. However, they are independent of the channel and transceiver, but they are dependent on the configuration of the FEC code. Hence the LUTs T_e and T_p have to be regenerated, when the FEC is reconfigured. *Note that all the LUTs are independent of the video sequences and that these LUTs are generated during the design process.* The size of these LUTs depends on

⁸The value of r_0 ranges from 0.2 to 1 with a constant step of 0.05.

the variables, hence the LUTs T_b , T_e , T_p have sizes of n_{snr} , $(n_{snr} \cdot n_I \cdot n_r)$, $(n_{snr} \cdot n_I \cdot n_r)$, respectively.

3) *Estimation of $p(\cdot)$* :: The CND and VND are involved in the PER estimation process, as well as in the search through the LUTs. As detailed in Section III-A, the MI flows of the VND and CND impose a low computational complexity. Furthermore, in our system the values of each variable are chosen using a constant step-size, which guarantees low complexity, while searching through the LUTs. For example, let us assume that the vector $[0 : 0.01 : 1]$ is employed for representing the values of I_s for generating the LUTs. In this case, the I_s value of 0.5 can be directly located at index 50.

4) *Full Search of r_0* :: n_r coding rates of r_0 may be tentatively tested, while r_1 is determined by Eq. (6) for a given r_0 . Hence the complexity of the full search may be expressed as $O(C \times n_r)$, where C is a constant, representing the complexity of estimating $p(L_i)$. However, when more layers are encapsulated in the source bitstream, the full-search based complexity increases exponentially, leading to a multi-dimensional optimization problem, which has been widely studied in the literature [45]–[47]. Specifically, the adaptive particle swarm optimization (APSO) technique of [47] may be readily employed for finding the global optimum in real-time. Note that in the scenarios where as few as 2–4 layers in the range of 2 to 4 are generated, even elite-force full-search is realistic at a modest complexity.

5) *Delay*:: Our technique is implemented using the SVC of H.264 as detailed in Section IV, where each video frame can be encoded into a single BL and multiple ELs. Since the IL encoding and decoding process is performed within each frame, no extra delay is imposed by our proposed technique. Hence, the system can be applied in low-delay applications.

IV. SYSTEM PERFORMANCE

In this section, we benchmark our proposed Optimized-IL-RSC system against an identical RSC-aided system employing the traditional UEP technique. The generator polynomials of [1011, 1101, 1111] are employed for configuring the RSC codec. Moreover, BPSK modulated signals were transmitted through *non-dispersive uncorrelated Rayleigh fading wireless channels*. The 30-frame Foreman and Football video clips represented in (176×144) -pixel quarter common intermediate format (QCIF) and 4 : 2 : 0 YUV format were encoded by the JSVM H.264/AVC reference video codec. The Football and Foreman sequences exhibit different motion-activity, hence they allow us to demonstrate the universal nature of our proposed system. These sequences were scanned at 15 and 30 FPS, respectively. The “frame-copy” based error concealment tool built into the JSVM H.264/AVC reference codec was activated for combating the effects of channel impairments. The GOP duration was set to 15, hence an Intra-frame-coded (I) picture was inserted every 15 frames. Correspondingly, both of the two video sequences were encoded into GOPs, consisting of an I frame, followed by 14 predicted (P) frames. Since the bi-directionally predicted (B) frames may impose error propagation on their forward- and backward-predicted dependent-frames, the B frames are disabled in our JSVM configuration. Additionally, only the Medium Grain Scalability (MGS) [4], [48] feature is

TABLE I
THE PARAMETERS OF THE VIDEO SEQUENCES EMPLOYED

	Football	Foreman
Representation	YUV 4:2:0	YUV 4:2:0
Format	QCIF	QCIF
Bits Per Pixel	8	8
FPS	15	30
Number of Frames	30	30
Video Codec	SVC-H.264	SVC-H.264
Bitrate	2297 kbps	218 kbps
Error-Free PSNR	38.8 dB	37.3 dB
Error Concealment	Frame-Copy	Frame-Copy

enabled, when encoding the video sequences into three different-quality ELs, namely into the layers L_0 , L_1 and L_2 using the standardized quantization parameters (QP) of 40, 32 and 24, respectively. Furthermore, each frame will be encoded into a single BL and two ELs, resulting in three network abstraction layer units (NALUs). These configurations jointly result in a bitrate of 2297 kbps and a PSNR of 38.8 dB for the Football sequence in the absence of transmission errors, while achieving 37.3 dB PSNR at 218 kbps for the Foreman sequence. The system parameters are listed in Table I.

Moreover, each SVC H.264/AVC-compressed bitstream was channel encoded and transmitted on a NALU by NALU [5] basis, which is the smallest unit to be decoded by the SVC decoder. Each NALU was protected by CRC codes. At the receiver, each decoded NALU failing to pass the CRC check process was removed before the SVC video decoding process. In all of our experiments, the compressed bitstreams were transmitted 300 times in order to generate statistically sound performance curves.

A. Off-line LUTs Generation

In our experiments, the vectors of $[0 : 0.5 : 15]$, $[0 : 0.01 : 1]$, $[0.33 : 0.02 : 1]$ ⁹ are utilized for the variables SNR , I_s , r , respectively, for generating the LUTs, which result in $n_{snr} = 31$, $n_I = 101$, $n_r = 33$. Firstly, we recorded the MI values of the LLRs output by the BPSK demodulator for the SNR values of $[0 : 0.5 : 15]$ for the sake of generating the LUT $T_b(SNR)$. Then for each SNR , the MI values $[0 : 0.01 : 1]$ for I_s and the coding rates $[0.33 : 0.02 : 1]$ for r were utilized for generating the LUTs $T_e(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$, simultaneously. Furthermore, 8-byte floating values were utilized for storing the LUTs in memory. Correspondingly, the LUTs T_b , T_e and T_p require memory sizes of about 248 bytes, 800 KB and 800 KB, respectively. Finally, $n_r = 33$ results in a full search complexity of $(C \times n_r^2 = C \cdot 1089)$, thereby finding the optimized coding

⁹These values can be stored as floats in 8 bytes each.

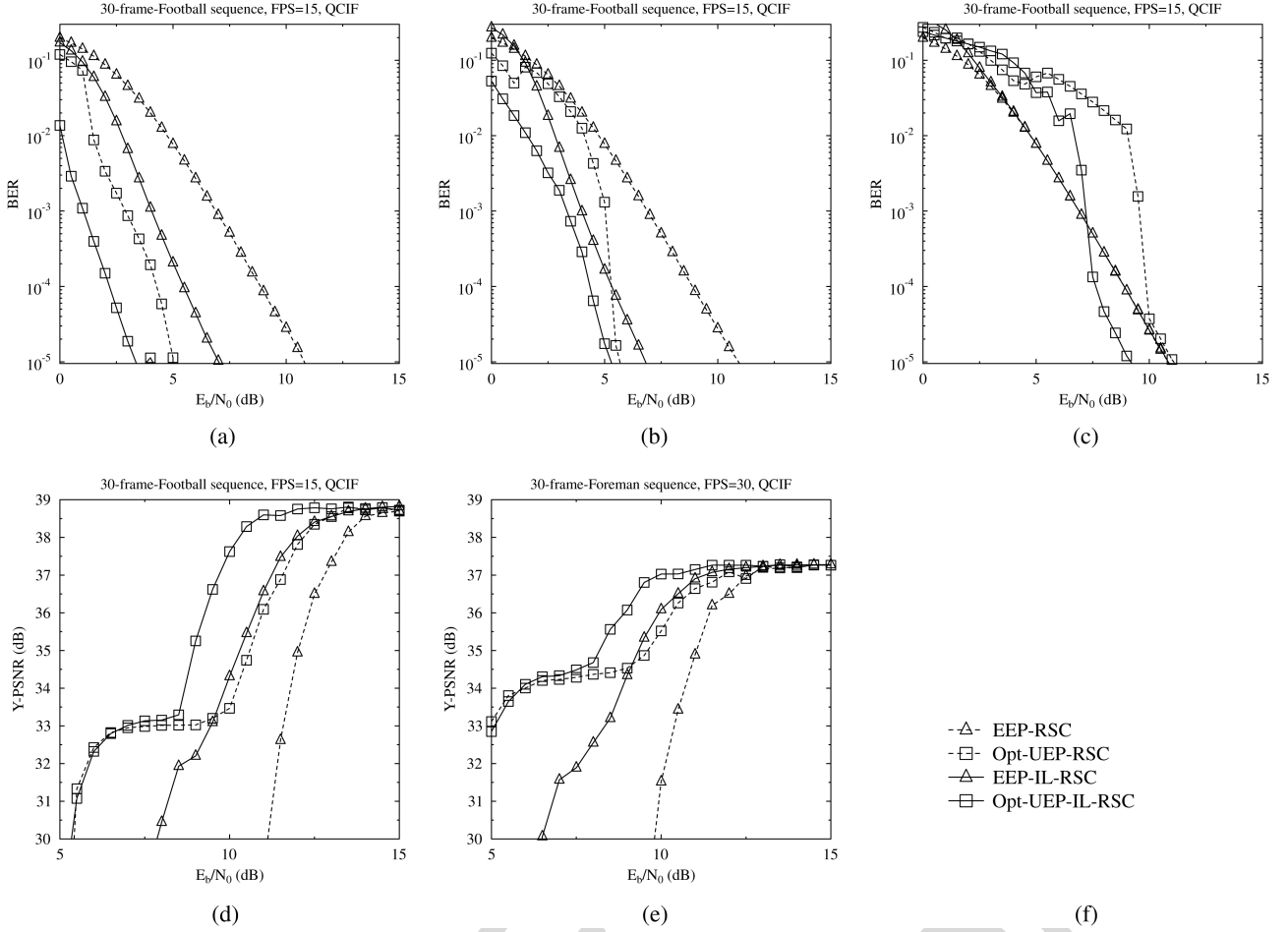


Fig. 8. BER and PSNR versus E_b/N_0 performance comparison of the proposed system and the benchmarks, namely the EEP-RSC scheme, the Opt-UEP-RSC scheme [20], the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme for the *Football* and *Foreman* sequences. *Non-dispersive uncorrelated Rayleigh fading wireless channels*. (a) BER vs E_b/N_0 for L_0 -Football. (b) BER vs E_b/N_0 for L_1 -Football. (c) BER vs E_b/N_0 for L_2 -Football. (d) PSNR vs E_b/N_0 for Football. (e) PSNR vs E_b/N_0 for Foreman. (f) Tested schemes.

TABLE II
EXAMPLE OF THE LUTs (a) $T_b(SNR)$.
(b) $T_e(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$

SNR	I_s	SNR	I_a	r	I_e	$p(l)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
1	0.45	3	0.67	0.87	0.81	0.93
2	0.50	3	0.67	0.88	0.83	0.89
3	0.56	3	0.67	0.89	0.85	0.84
4	0.62	3	0.67	0.90	0.87	0.77
5	0.67	3	0.67	0.91	0.89	0.69
6	0.72	3	0.67	0.92	0.91	0.57
7	0.76	3	0.67	0.93	0.93	0.46
8	0.80	3	0.67	0.94	0.94	0.33
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

(a) $T_b(SNR)$ (b) $T_e(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$

rates of r_0 , r_1 and r_2 . A number of entries of the LUTs generated for our system are displayed in Table II.

B. System Performance

In this section, we benchmark our system against the traditional equal error protection (EEP) system, which is referred to

as EEP-RSC. Furthermore, the traditional optimal UEP-RSC (Opt-UEP-RSC) system is also presented, which is the system of [20] applied in the scenario of an RSC coded system, where all the layers are unequally protected by an RSC code dispensing with the IL technique. Moreover, the IL aided EEP system (EEP-IL-RSC) is also considered.

The BER curves of the L_0 , L_1 and L_2 layers of the Football sequence are displayed in Figs. 8a, 8b and 8c, respectively, where the Opt-UEP-IL-RSC system always outperforms the Opt-UEP-RSC system. Moreover, as seen in Fig. 8c, the EEP schemes outperform the Opt-UEP-IL-RSC and Opt-UEP-RSC schemes in the lower E_b/N_0 range. This is due to the fact that the coding-rate of layer L_2 is sacrificed for the sake of protecting the L_0 and L_1 layers. Similar trends were observed for the Foreman sequence, which are not included here owing to space-economy.

The PSNR versus E_b/N_0 performance recorded for the Football sequence is displayed in Fig. 8d, where we observe that the EEP-IL-RSC scheme substantially outperforms the EEP-RSC system, while it only slightly outperforms the Opt-UEP-RSC scheme. However, by optimizing the coding rates, the Opt-UEP-IL-RSC scheme becomes capable of substantially outperforming the Opt-UEP-RSC system. In the low E_b/N_0 range, the Opt-UEP-IL-RSC and Opt-UEP-RSC



Fig. 9. Comparison of frames at E_b/N_0 of 10 dB for the *Football* sequences. The five columns (from left to right) indicate frames of the original video, the EEP-RSC scheme, the Opt-UEP-RSC [20] scheme, the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme, respectively.

schemes have a similar PSNR performance and tend to exhibit a residual error floor. This is due to the fact that these two systems endeavor to protect the BL L_0 by sacrificing the protection of the ELs, where the ELs are unlikely to be recovered before reaching $E_b/N_0 = 10$ dB. This error floor becomes explicit in the BER curves of L_2 displayed in Fig. 8c. On the other hand, the Opt-UEP-RSC scheme outperforms the EEP-RSC scheme by about $E_b/N_0 = 1.3$ dB at a PSNR of 38 dB. The Opt-UEP-IL-RSC scheme achieves an E_b/N_0 reduction of about 1.9 dB compared to the Opt-UEP-RSC scheme at a PSNR of 38 dB. Alternatively, about 3.3 dB of PSNR video quality improvement is observed at an E_b/N_0 of 10 dB. A subjective comparison of the benchmarks recorded for the *Football* sequence is presented in Fig. 9.

For providing further insights for video scenes having different motion-activity, the PSNR versus E_b/N_0 performance of the Opt-UEP-IL-RSC is portrayed in Fig. 8e using the *Foreman* sequence. The Opt-UEP-RSC scheme outperforms the EEP-RSC system by about $E_b/N_0 = 0.8$ dB at a PSNR of 37 dB. Moreover, about 1.7 dB of power reduction is achieved by the Opt-UEP-IL-RSC scheme compared to the Opt-UEP-RSC scheme at a PSNR of 37 dB. Alternatively, about 1.6 dB of PSNR video quality improvement may be observed at an E_b/N_0 of 10 dB.

C. Optimized Coding Rates

The optimized coding rates found by our proposed regime for the *Football* sequence are shown in Fig. 10 in comparison to the benchmarks. Specifically, the y axis of Fig. 10 indicates the averaged coding rates of all frames. Furthermore, the coding rates of L_0 and L_2 are presented, while the curves of L_1 are similar to those of the BL L_0 . Observe in Fig. 10 that in the low E_b/N_0 range the coding rates of the BL L_0 of the Opt-UEP-RSC and Opt-UEP-IL-RSC schemes are lower than the overall coding rate of 0.5. The reason of this observation is that the protection of the layers L_1 and L_2 is sacrificed for the sake of protecting the more important BL L_0 . At high E_b/N_0 values, the coding rates of the BL L_0 are increased due to the fact that L_0 is more likely to be corrected at high coding rates, when favorable channel conditions prevail. In comparison to the BL L_0 , the coding rates of L_2 are reduced upon increasing E_b/N_0 .

V. CONCLUSIONS

We proposed an optimization technique for bit-level IL-FEC coded scalable video transmission over wireless channels,

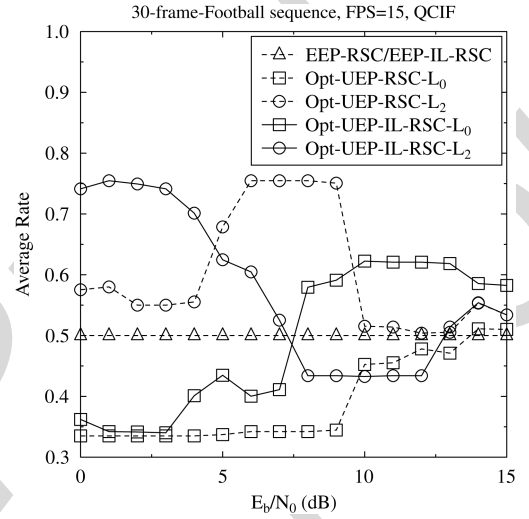


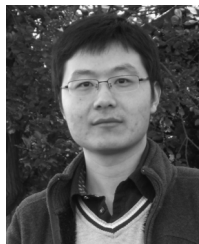
Fig. 10. Coding rates comparison of the benchmarks, namely the EEP-RSC scheme, the Opt-UEP-RSC [20] scheme, the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme, for the *Football* sequence. *Non-dispersive uncorrelated Rayleigh fading wireless channels*

where soft-decoding aided FEC is utilized. Firstly, LUTs were generated for the sake of optimizing the wireless transceivers and the FEC codec, which relied on exploiting the off-line training based MI. Moreover, the MI exchange of the VND and CND of Fig. 2 was exploited for realistic video systems, where both the PERs of the IL-FEC coded BL and of the FEC coded ELs were estimated, based on the knowledgeable specific coding rates of all the scalable video layers. Then the expected PSNR distortion was minimized by performing a full search. BPSK modulation, an RSC FEC code and the SVC-H.264 video codec were employed in our design example, but our solution may be readily applied for arbitrary transceivers generating near-Gaussian distributed LLRs, for arbitrary non-iteratively decoded channel codes and for arbitrary layered video transmissions. Our simulation results show that the proposed optimized IL-FEC system outperforms the traditional optimal UEP system by about 1.9 dB of E_b/N_0 or 3.3 dB of PSNR, when employing a RSC code.

In our future work, we will incorporate the IL-FEC scheme into multiview video coding and we will also carry out further investigations for applying our solution to more advanced FEC codes, such as turbo codes [34], LDPC codes, BICM, BICM-ID, TCM aided TCM schemes.

REFERENCES

- [1] L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG2, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. New York, NY, USA: Wiley, 2007.
- [2] T. Zhang and Y. Xu, "Unequal packet loss protection for layered video transmission," *IEEE Trans. Broadcast.*, vol. 45, pp. 243–252, Jun. 1999.
- [3] H. Imaizumi and A. Luthra, "Three-Dimensional Television, Video and Display Technologies," in *ch. MPEG-2 Multiview Profile*. Berlin, Germany: Springer Verlag, 2002, pp. 169–181.
- [4] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, pp. 1103–1120, Sep. 2007.
- [5] "Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG," ITU-T Rec. H.264/ISO/IEC 14496-10 AVC: Advanced Video Coding for Generic Audiovisual Services, Mar. 2010.
- [6] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proc. IEEE*, vol. 99, pp. 626–642, Apr. 2011.
- [7] F. Yang, Q. Zhang, W. Zhu, and Y.-Q. Zhang, "End-to-end TCP-friendly streaming protocol and bit allocation for scalable video over wireless Internet," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 777–790, May 2004.
- [8] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Trans. Inf. Theory*, vol. 13, pp. 600–607, Oct. 1967.
- [9] T. Brüggen and P. Vary, "Unequal error protection by modulation with unequal power allocation," *IEEE Commun. Lett.*, vol. 9, pp. 484–486, Jun. 2005.
- [10] V. Pavlushkov, R. Johannesson, and V. Zyablov, "Unequal error protection for convolutional codes," *IEEE Trans. Inf. Theory*, vol. 52, pp. 700–708, Feb. 2006.
- [11] J. Hagenauer, "Rate-compatible puncture convolutional codes (RCPC) and their application," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, Apr. 1988.
- [12] N. Rahnavard and F. Fekri, "New results on unequal error protection using LDPC codes," *IEEE Commun. Lett.*, vol. 10, pp. 43–45, Jan. 2006.
- [13] V. Kumar and O. Milenkovic, "On unequal error protection LDPC codes based on Plotkin-type constructions," *IEEE Trans. Commun.*, vol. 54, pp. 994–1005, Jun. 2006.
- [14] C. Gong, G. Yue, and X. Wang, "Message-wise unequal error protection based on low-density parity-check codes," *IEEE Trans. Commun.*, vol. 59, pp. 1019–1030, Apr. 2011.
- [15] N. Rahnavard, H. Pishro-Nik, and F. Fekri, "Unequal error protection using partially regular LDPC codes," *IEEE Trans. Commun.*, vol. 55, pp. 387–391, Mar. 2007.
- [16] 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.
- [17] S. X. Ng, J. Y. Chung, and L. Hanzo, "Turbo-detected unequal protection MPEG-4 wireless video telephony using multi-level coding, trellis coded modulation and space-time trellis coding," *IEE Proc. Commun.*, vol. 152, pp. 1116–1124, Dec. 2005.
- [18] H. Wang, F. Zhai, Y. Eisenberg, and A. Katsaggelos, "Cost-distortion optimized unequal error protection for object-based video communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 1505–1516, Dec. 2005.
- [19] M. Aydinlik and M. Salehi, "Turbo coded modulation for unequal error protection," *IEEE Trans. Commun.*, vol. 56, pp. 555–564, Apr. 2008.
- [20] H. Ha and C. Yim, "Layer-weighted unequal error protection for scalable video coding extension of H.264/AVC," *IEEE Trans. Consumer Electron.*, vol. 54, pp. 736–744, May 2008.
- [21] Y. C. Chang, S. W. Lee, and R. Komiya, "A fast forward error correction allocation algorithm for unequal error protection of video transmission over wireless channels," *IEEE Trans. Consumer Electron.*, vol. 54, pp. 1066–1073, Aug. 2008.
- [22] Y. C. Chang, S. W. Lee, and R. Komiya, "A low complexity hierarchical QAM symbol bits allocation algorithm for unequal error protection of wireless video transmission," *IEEE Trans. Consumer Electron.*, vol. 55, pp. 1089–1097, Aug. 2009.
- [23] D. Sejdinović, D. Vukobratović, A. Doufexi, V. Šenk, and R. Piechocki, "Expanding window fountain codes for unequal error protection," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2510–2516, 2009.
- [24] D. Vukobratović, V. Stanković, D. Sejdinović, L. Stanković, and Z. Xiong, "Scalable video multicast using expanding window fountain codes," *IEEE Trans. Multimedia*, vol. 11, no. 6, pp. 1094–1104, 2009.
- [25] E. Maani and A. Katsaggelos, "Unequal error protection for robust streaming of scalable video over packet lossy networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, pp. 407–416, Mar. 2010.
- [26] S. Ahmad, R. Hamzaoui, and M. Al-Akaidi, "Unequal error protection using fountain codes with applications to video communication," *IEEE Trans. Multimedia*, vol. 13, pp. 92–101, Feb. 2011.
- [27] K. Nguyen, T. Nguyen, and S.-C. Cheung, "Video streaming with network coding," *J. Signal Process. Syst.*, vol. 59, pp. 319–333, Jun. 2010.
- [28] C. Hellge, D. Gomez-Barquero, T. Schierl, and T. Wiegand, "Layer-aware forward error correction for mobile broadcast of layered media," *IEEE Trans. Multimedia*, vol. 13, pp. 551–562, Jun. 2011.
- [29] Nasruminallah and L. Hanzo, "Near-capacity H.264 multimedia communications using iterative joint source-channel decoding," *IEEE Commun. Surveys and Tutorials*, vol. 14, pp. 538–564, 2012, Second Quarter.
- [30] M. Hallouash and H. Radha, "Network coding with multi-generation mixing: A generalized framework for practical network coding," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 466–473, 2011.
- [31] Y. Huo, M. El-Hajjar, and L. Hanzo, "Inter-layer FEC aided unequal error protection for multilayer video transmission in mobile TV," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1622–1634, 2013.
- [32] D. Wu, Y. T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the Internet: Challenges and approaches," *Proc. IEEE*, vol. 88, no. 12, pp. 1855–1877, 2000.
- [33] T. Stockhammer, M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657–673, 2003.
- [34] L. Hanzo, T. Liew, B. Yeap, and R. Tee, *Turbo Coding, Turbo Equalisation and Space-Time Coding*. New York, NY, USA: Wiley, 2011.
- [35] T. Stockhammer, A. Shokrollahi, M. Watson, M. Luby, and T. Gasiba, "Application Layer Forward Error Correction for Mobile Multimedia Broadcasting," in *Handbook of Mobile Broadcasting: DVB-H, DMB, ISDB-T and Media FLO*, B. Furhet and S. Ahson, Eds. Boca Raton, FL, USA: CRC, 2008, pp. 239–280.
- [36] C. Hellge, T. Schierl, and T. Wiegand, "Multidimensional layered forward error correction using rateless codes," in *Proc. IEEE Int. Conf. Communications*, Beijing, China, May 2008, pp. 480–484.
- [37] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, pp. 670–678, Apr. 2004.
- [38] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, pp. 429–445, Mar. 1996.
- [39] J. Chen, A. Dholakia, E. Eleftheriou, M. Fossorier, and X.-Y. Hu, "Reduced-complexity decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 53, pp. 1288–1299, Aug. 2005.
- [40] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE Int. Conf. Communications*, Geneva, Switzerland, May 1993, pp. 1064–1070.
- [41] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, Oct. 2001.
- [42] R. Maunder and L. Hanzo, "Extrinsic information transfer analysis and design of block-based intermediate codes," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 762–770, Mar. 2011.
- [43] R. Khalili and K. Salamati, "A new analytic approach to evaluation of packet error rate in wireless networks," in *Proc. Communication Networks and Services Research Conf.*, Halifax, NS, Canada, May 2005, pp. 333–338.
- [44] X. Yang, C. Zhu, Z. Li, X. Lin, G. Feng, S. Wu, and N. Ling, "Unequal loss protection for robust transmission of motion compensated video over the Internet," *Signal Process.: Image Commun.*, vol. 18, no. 3, pp. 157–167, 2003.
- [45] M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Computat.*, vol. 6, no. 1, pp. 58–73, 2002.
- [46] D. Schonfeld and N. Bouaynaya, "A new method for multidimensional optimization and its application in image and video processing," *IEEE Signal Process. Lett.*, vol. 13, no. 8, pp. 485–488, 2006.
- [47] Z.-H. Zhan, J. Zhang, Y. Li, and H.-H. Chung, "Adaptive particle swarm optimization," *IEEE Trans. Syst., Man, Cybern. B: Cybern.*, vol. 39, no. 6, pp. 1362–1381, 2009.
- [48] A. Detti, G. Bianchi, C. Pisa, F. Proto, P. Loret, W. Kellerer, S. Thakolsri, and J. Widmer, "SVEF: An open-source experimental evaluation framework for H.264 scalable video streaming," in *Proc. IEEE Symp. Computers and Communications, ISCC*, Sousse, Tunisia, Jul. 2009, pp. 36–41.



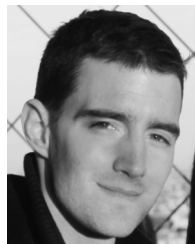
Yongkai Huo received the B.Eng. degree with distinction in computer science and technology from Hefei University of Technology, Hefei, China, in 2006 and the M.Eng. degree in computer software and theory from University of Science and Technology of China, Hefei, China, in 2009. He is currently working toward the Ph.D. degree with the Communications, Signal Processing and Control Group, School of Electronics and Computer Science, University of Southampton, Southampton, U.K. He received a scholarship under the China-U.K. Scholarships for Excellence Programme.

His research interests include distributed video coding, multiview video coding, robust wireless video streaming and joint source-channel decoding.



Mohammed El-Hajjar is a lecturer in the Electronics and Computer Science in the University of Southampton. He received his BEng degree in Electrical Engineering from the American University of Beirut, Lebanon in 2004. He then received an MSc in Radio Frequency Communication Systems and PhD in Wireless Communications both from the University of Southampton, UK in 2005 and 2008, respectively. Following the PhD, he joined Imagination Technologies as a research engineer, where he worked on designing and developing the

BICM peripherals in Imagination's multi-standard communications platform, which resulted in several patent applications. In January 2012, he joined the Electronics and Computer Science in the University of Southampton as a lecturer in the Communications, Signal Processing and Control research group. He is the recipient of several academic awards and has published a Wiley-IEEE book and in excess of 30 journal and international conference papers. His research interests are mainly in the development of intelligent communications systems for the Internet of Things including massive MIMO systems for mm-wave communications, cooperative communications and Radio over fibre systems.



Robert G. Maunder has studied with Electronics and Computer Science, University of Southampton, UK, since October 2000. He was awarded a first class honors BEng in Electronic Engineering in July 2003, as well as a PhD in Wireless Communications and a lectureship in December 2007. His research interests include joint source/channel coding, iterative decoding, irregular coding and modulation techniques. He has published a number of IEEE papers in these areas.



Lajos Hanzo (F'04) received the degree in electronics in 1976, the Doctorate degree in 1983, and the D.Sc. degree.

During his 35-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 80 PhD students, co-authored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1300 research entries at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European IST Programme and the Mobile Virtual Centre of Excellence (VCE), UK. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and since 2009 he has been a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to <http://www-mobile.eecs.soton.ac.uk>.

Layered Wireless Video Relying on Minimum-Distortion Inter-Layer FEC Coding

Yongkai Huo, Mohammed El-Hajjar, Robert G. Maunder, and Lajos Hanzo, *Fellow, IEEE*

Abstract—Layered video coding is capable of progressively refining the reconstructed video quality with the aid of multiple layers of unequal importance. When the base layer (BL) is corrupted or lost due to channel impairments, the enhancement layers (ELs) must be discarded by the video decoder, regardless whether they are perfectly decoded or not, which implies that the transmission power assigned to the ELs is wasted. To circumvent this problem, we proposed a bit-level inter-layer forward error correction (IL-FEC) scheme for layered video transmission in our previous work, which implanted the systematic bits of the BL into the systematic bits of the ELs using exclusive-OR operations (XOR). This allowed the receiver to exploit the implanted bits of the ELs for assisting the BL's decoding and hence improved the overall system performance of our IL-FEC aided layered video scheme. In this treatise, we find the specific FEC coding rates in a real-time on-line fashion for the sake optimizing the overall system performance. The proposed procedure is widely applicable to diverse wireless transceivers and FEC codecs. Our simulation results show that the proposed optimized IL-FEC system outperforms the traditional optimal UEP by about 1.9 dB of E_b/N_0 at a peak signal-to-noise ratio (PSNR) of 38 dB. Viewing the improvements in terms of the video quality, 3.3 dB of PSNR improvement is attained at an E_b/N_0 of 10 dB, when employing a recursive systematic convolutional (RSC) code.

Index Terms— *[Author, please supply index terms/keywords for your paper. To download the IEEE Taxonomy go to http://www.ieee.org/documents/2009Taxonomy_v101.pdf].*

I. INTRODUCTION

IN RECENT years, low-delay lip-synchronized wireless video communications became a reality and the road to this is detailed in [1]. Layered video coding [2] has been widely adopted by the existing video coding standards [3]–[6], which is capable of generating multiple layers of unequal importance. The most important layers are referred to as the base layer (BL), while the less important layers depend on the BL and are

referred to as enhancement layers (ELs). As an inherent component of layered video coding standards, partitioned mode video coding [5] encodes each video-frame slice into at most three partitions, namely partition A, B and C according to the error sensitivity of the bitstreams, where partition A may be deemed to be the BL, while partitions B and C are treated as ELs. Furthermore, the moving picture expert group (MPEG)'s [1] developed a multiview profile (MVP) [3], where the left view and right view are encoded as a BL and an EL, respectively. Another recent layered video coding standard, referred to as scalable video coding (SVC) [4], [5] was designed as an extension of H.264/AVC [5]. In SVC, a video sequence is encoded into a stream containing multiple dependent layers, where a number of subset layers may be extracted and hence the bitrate may be controlled according to the users' specific requirements in practical video services. A common characteristic of layered video coding scheme is that the ELs may be assigned a lower priority than the BL and hence may be discarded in the scenario of network congestion or buffer overflow [7]. Moreover, when the related BL is corrupted or lost due to channel impairments, the layered video decoder has to discard the corresponding ELs, regardless whether they are correctly recovered or not.

Unequal error protection (UEP) was first proposed in [8], where a stronger forward error correction (FEC) code was allocated to the more important data and a weaker FEC code was allocated to the less important data. In [9] UEP was achieved by optimizing the transmission power assigned to individual source bits according to their importance. Additionally, UEP using convolutional codes was investigated in [10], where the rate-compatible convolutional codes (RCPC) proposed in [11] were employed. Moreover, a number of UEP techniques based on low-density parity-check (LDPC) codes were studied in [12]–[15], where the more important bits were more strongly protected than the less important bits. Moreover, most of the above-mentioned UEP contributions considered artificial signals, rather than realistic video signals, which may not be directly extended to the scenario of flawless, low-delay, lip-synchronized video communications.

In layered video coded bitstreams different layers have a different importance, which implies that employing UEP techniques would substantially improve the reconstructed video quality. Generally, these UEP techniques [16]–[31] may be classified into two categories,¹ namely the packet-level schemes [18], [20], [23]–[27], [30] and bit-level schemes [16], [17], [19], [21], [22], [28], [29], [31]. Specifically, the packet-level contributions [18], [20], [23]–[27], [30] mitigate the packet loss

Manuscript received June 02, 2013; revised November 04, 2013; accepted November 30, 2013. This work was supported by the EU's Concerto project, of the EPSRC under the auspices of the India-UK Advanced Technology Centre (IU-ATC) and that of the ERC's Advanced Fellow Grant. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Wenwu Zhu.

The authors are with the School of Electronics and Computer Science, University of Southampton, U.K. (e-mail: yh3g09@ecs.soton.ac.uk; meh@ecs.soton.ac.uk; rm@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2014.2300449

¹Thanks to an anonymous reviewer for this valuable suggestion.

events as exemplified by the packets lost in Internet-routers [32], while the bit-level ones are devoted to eliminating bit-errors of wireless scenarios [33]. Let us continue by considering the family of packet-level UEP schemes first [18], [20], [25], [26]. An UEP scheme was conceived for object-based video communications in [18] for achieving the best attainable system performance under specific bitrate and delay constraints in an error-prone network environment. A novel UEP method using Reed-Solomon codes was proposed in [20] for SVC video transmission over networks inflicting packet-loss events. Firstly, the layer-weighted expected zone of error propagation (LW-EZEP) was defined as an efficient performance metric for quantifying the error propagation effects imposed by packet loss events. Then, the corresponding Reed-Solomon (RS) coding rates were assigned based on LW-EZEP for minimizing the expected video distortion. Cross-layer operation aided scalable video streaming designed for error-prone channels was investigated in [25], where the RS coded UEP was optimized for robust video delivery. The expected video quality was evaluated based on both the available bandwidth and the packet loss ratio (PLR) encountered, which was then further improved by employing content-aware bitrate allocation. Finally, an error concealment method was invoked at the receiver. The authors of [26] studied an UEP scheme using Luby Transform (LT) codes for recovering the video packets lost owing to network congestions. The above-mentioned UEP methodologies tended to employ hard decoded FEC [34] codes and dispensed with soft-information exchange in the decoding process, since they are carried out at the application layer. Similarly, bit-level UEP techniques have also been widely investigated [16], [17], [19], [21], [22], [29]. In [17], an UEP scheme using a turbo transceiver was optimized for wireless video telephony. The UEP performance of data-partitioned [5] H.264/AVC video streaming systems using recursive systematic convolutional (RSC) codes was evaluated in [29], while turbo coded modulation based UEP was investigated in [19], where both the cutoff rates and the channel capacity of each of the UEP levels was determined. The authors of [21] considered the unequal importance of both the video-frames in a group of pictures (GOP) and the significance of the diverse macroblocks (MB) in a video frame for transmission over wireless channels, where a prompt and efficient fast rate allocation scheme was also investigated. However, only three protection classes were discussed in [21], which limits the attainable system performance. The same authors also considered the different importance of the intra-coded (I) frame and of the predicted (P) frames within a GOP in [22], where the video bits of different importance were mapped to the different-integrity modulation constellation points of hierarchical quadrature amplitude modulation (QAM). These bit-level schemes tended to employ physical layer FEC codes [35] and perform soft decoding [34] for wireless video communications.

In the traditional UEP scheme conceived for layered video communication, variable-rate FEC codes were assigned to the different-sensitivity layers for improving the reconstructed video quality. However, when the BL is corrupted or lost, the ELs have to be discarded by the video decoder, regardless whether they are perfectly decoded or not, which implies that

both the transmission power and the bandwidth assigned to the ELs is wasted. Hence it is beneficial improve the protection of the more important BL with the aid of the ELs. Hence, the authors of [24] proposed a packet-level solution based on fountain codes [23] for scalable video multicast over networks inflicting packet loss events, where ELs conveyed parity information protecting the more important BL. By contrast, hierarchical network coding [27] and generalized network coding [30] techniques were also proposed for combating packet-loss events, when transmitting information of unequal importance. Furthermore, a bit-level layer-aware FEC (LA-FEC) philosophy using a hard-decoded Raptor code was designed for scalable video transmission over the binary erasure channels (BEC) in [28], [36]. The Raptor encoder generated the parity bits right across the BL and the ELs at the transmitter. As a benefit, the parity bits of the ELs may be utilized for assisting in correcting the errors residing in the BL at the receiver. Motivated by these advances, we developed a bit-level inter-layer coded FEC (IL-FEC) scheme layered video telephony over wireless fading channels relying on a soft-decoded systematic FEC code in [31], where the systematic bits of the BL are implanted into the ELs at the transmitter. At the receiver, the BL's bits implanted into the ELs may be utilized for correcting the BL. The above-mentioned IL-FEC technique of [31] was also combined with the UEP philosophy for the sake of further improving the attainable system performance. In [31], a number of coding rates were tested for the sake of proving the benefits of the proposed IL technique, where the code rates arrangements were determined empirically. However, in practical scenarios, different configurations of video codecs and different video sequences may have different characteristics, which may require different channel coding rates for achieving the best system performance.

Against this background, in this treatise, we propose a technique for finding the optimized coding rates for coded bitstreams "on-the-fly" at the transmitter, which optimizes the IL-FEC coded system performance. Specifically, we find the coding rates achieving the minimum video quality distortion with the aid of the mutual information (MI) between the log-likelihood ratios (LLRs) and the corresponding video bits. In this context, the soft-decoding metric of the FEC codec and of the demodulator are characterized by lookup tables (LUTs), since these cannot be characterized theoretically. Finally, the widely employed SVC-H.264 codec will be utilized for the sake of benchmarking our proposed system. In a nutshell, we focus our attention on the optimization of bit-level IL-FEC encoded scalable video communications over wireless fading channels. Specifically, our proposed technique is significantly different from the existing contributions [16], [20], [21], [23], [25], [26], [28] in the following aspects. Firstly, our technique is proposed for layered video communication over wireless channels, while the packet-level contributions of [20], [23], [25], [26] consider networks, which inflict packet-loss events or on networks modeled by the BEC. Secondly, IL-FEC relies on bit-level soft-decoding aided channel decoders, such as an RSC code, while [20], [23], [25], [26], [28] considered hard-decoded Reed-Solomon or fountain codes. Moreover, we propose a solution employing near-continuous coding rates, while only

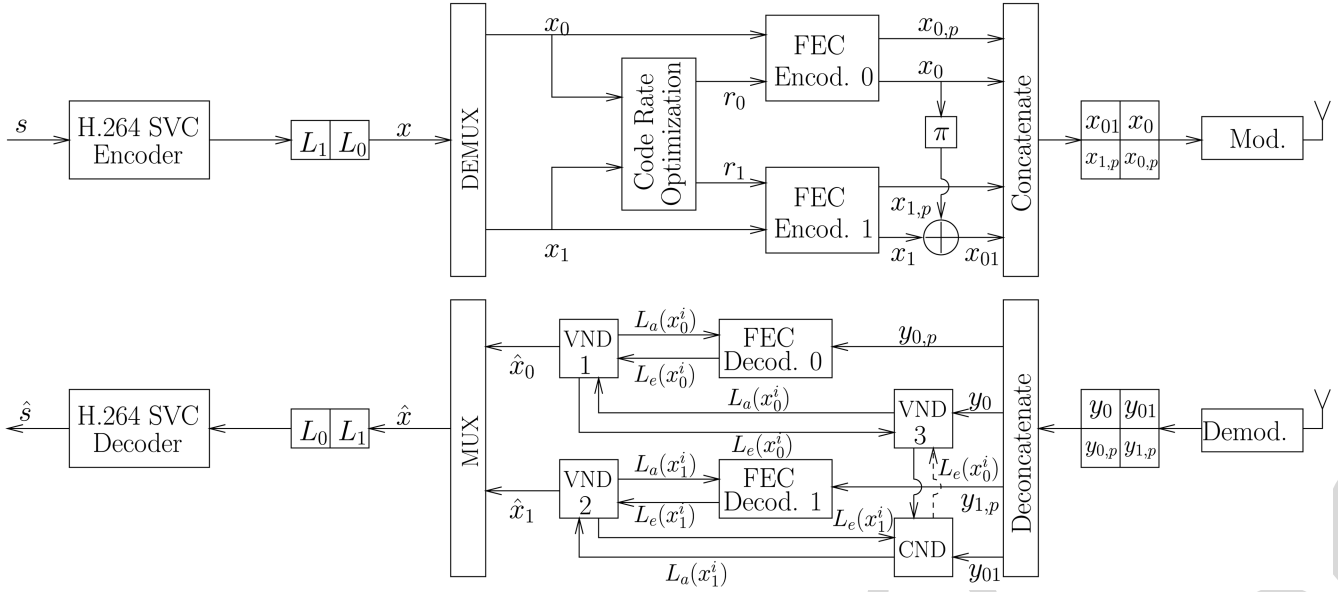


Fig. 1. Block diagram of the proposed IL-FEC coded SVC H.264/AVC coded video system, where r_0 and r_1 represent the code rates for FEC encoder 0 and 1, respectively. The “Code Rate Optimization” block will be detailed in Section III.

three protection rates - namely low-order, medium-order and high-order protection - were utilized in [16], [21], which limits the system’s performance. Finally, we find the optimized coding rates for the IL-FEC aided UEP system of [31], where the basic principle of IL-FEC was conceived, but the assisted focus and the system optimization was set aside for further research. Against this background, the rationale and novelty of this paper is summarized as follows.

- 1) *We set out to optimize soft-decoding bit-level IL-FEC encoded scalable video communication over wireless fading channels relying on a new objective function (OF) for the sake of finding the best code rate. Specifically, our technique relies on the MI and LUTs for finding the optimized coding rates in a real-time, online fashion, while minimizing video quality distortion. In contrast to the packet-level contributions of [20], [23], [25], [26], where the authors considered hard-decoded video streams transmitted over networks, which inflict packet-loss events or networks modelled by BEC, we consider transmission over non-dispersive Rayleigh fading wireless channels, where the layered video stream is protected by FEC codes and then soft-decoded at the receiver.*
- 2) *We utilize the MI and LUTs for the sake of optimizing the proposed system’s performance, where we show an E_b/N_0 gain of 1.9 dB at a PSNR of 38 dB. Alternatively, viewing the improvements in terms of the attainable video quality, or a PSNR gain of 3.3 dB is achieved at an E_b/N_0 of 10 dB over the benchmarker of [21].*
- 3) *A general solution is proposed for optimizing the coding rate allocation for arbitrary non-iterative FEC codes protecting, arbitrary layered video codecs. More explicitly, our solution is applicable to arbitrary transceivers provided that the receiver is capable of generating near-Gaussian distributed soft-information.*

We use the SVC-H.264/AVC scheme in our simulations, but our proposed scheme is not limited to SVC, it can be readily applied in any arbitrary system relying on layered video coding, such as the partitioned video coding regime of [31]. The rest of this paper is organized as follows. Section II details our proposed IL-FEC system model and the related video transmission techniques. Then the coding-rate optimization of the system is detailed in Section III. The performance of our optimized IL-FEC scheme using a RSC codec is benchmarked in Section IV using two video sequences having different motion characteristics. Finally, we offer our conclusions in Section V.

II. SYSTEM OVERVIEW

In this section, we briefly introduce the architecture of the inter-layer FEC scheme conceived for scalable video transmission, while further details on this technique may be found in [31]. The architecture of the system is displayed in Fig. 1, where the SVC mode of H.264 [5] is employed. Moreover, two layers, namely the BL L_0 and the EL L_1 are assumed to be generated by the SVC encoder, noting that more complex scenarios may be readily applied - we simply use two layers here for ease of explanation.² The variable node decoder (VND) and check node decoder (CND) [37] used in Fig. 1 are illustrated in Fig. 2. Both the VND and CND blocks accept a number of soft information inputs and generate the related soft information outputs for iteratively exploiting the deliberately imposed inter-dependencies between the coded layers L_0 and L_1 . Specifically, assuming that u_1 , u_2 and $u_3 = u_1 \oplus u_2$ are random binary variables, the VND of Fig. 2 sums two LLR inputs for generating an improved LLR output, which may be formulated as $L_{o3}(u_1) = L_{i1}(u_1) + L_{i2}(u_1)$. Given the confidence of the bits u_1 and u_2 , the boxplus operation of $L(u_3 = u_1 \oplus u_2) = L(u_1) \boxplus L(u_2)$

²In our results in Section IV, we show how our proposed techniques work for SVC encoder using three layers.

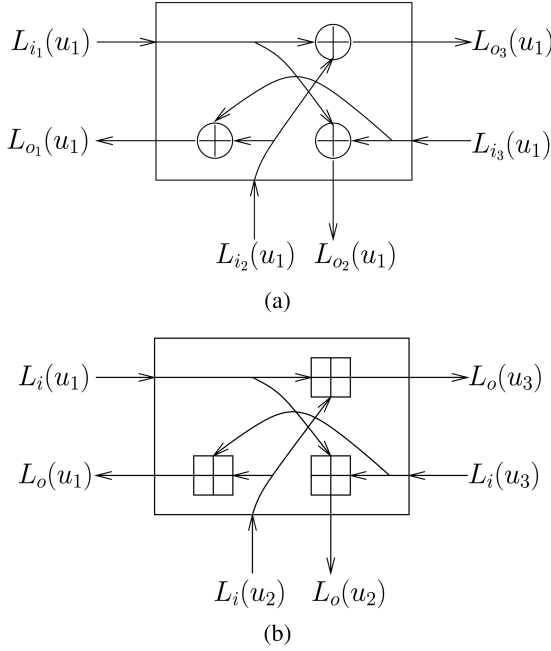


Fig. 2. The structure of VND and CND, where \oplus and \boxplus represent the addition and boxplus action, respectively. $L_i(\cdot)$ and $L_o(\cdot)$ represent the input and output LLR, respectively. (a) VND. (b) CND.

[38] may be utilized for deriving the confidence of the bit u_3 . Specifically, the boxplus operation \boxplus may be formulated as follows [39]

$$\begin{aligned} L(u_1) \boxplus L(u_2) &= \log \frac{1 + e^{L(u_1)} e^{L(u_2)}}{e^{L(u_1)} + e^{L(u_2)}} \\ &= \text{sign}[L(u_1)] \cdot \text{sign}[L(u_2)] \cdot \min[|L(u_1)|, |L(u_2)|] \\ &\quad + \log \left[1 + e^{-|L(u_1) + L(u_2)|} \right] \\ &\quad - \log \left[1 + e^{-|L(u_1) - L(u_2)|} \right]. \end{aligned} \quad (1)$$

In contrast to the VND block, the CND operation of Fig. 2 may be expressed as $L_o(u_3) = L_i(u_1) \boxplus L_i(u_2)$ for calculating the LLR of bit u_3 , given the LLR input of the bits u_1 and u_2 .

In Section II-A, the IL-FEC encoding techniques invoked at the transmitter are illustrated followed by our IL-FEC decoding techniques in Section II-B, where we emphasize on how the VND and the CND can exploit the IL information for improving the overall system performance. In Section II-A and II-B, we assume that the layers L_0 and L_1 contain an identical number of bits for simplifying our explanations, but our algorithm may be readily applied to the scenario of having an unequal number of bits as in [31].

A. Transmitter Model

At the transmitter of Fig. 1, the video source signal s is compressed using the SVC of the H.264 encoder, generating layers L_0 and L_1 . Then the output bitstream x is de-multiplexed into the bitstreams of x_0 and x_1 by the DEMUX block of Fig. 1, where x_0 and x_1 represent the bitstream of the BL L_0 and of the EL L_1 , respectively. Afterwards, the resultant two layers are encoded as follows:

- 1) The BL bit sequence x_0 representing L_0 will be encoded by the FEC encoder 0 of Fig. 1, which results in the encoded bits containing the systematic bits x_0 and parity bits $x_{0,p}$.
- 2) The bit sequence of the EL x_1 representing L_1 will firstly be encoded into the systematic bits x_1 and the parity bits $x_{1,p}$ by the FEC encoder 1. Then the XOR operation will be utilized for implanting the systematic information of x_0 into the systematic information of x_1 , which is carried out without changing the parity bits $x_{1,p}$ of the EL L_1 , as shown in Fig. 1. Specifically, the implantation process results in the check bits of $x_{01}^i = x_0^i \oplus x_1^i$. After this procedure, both the check bits x_{01}^i and the parity bits $x_{1,p}$ are output.

Finally, the bit sequences x_0 , $x_{0,p}$, x_{01} and $x_{1,p}$ are concatenated into a joint bitstream for transmission. Note however that the layers x_0 and x_1 may contain a different number of bits. Additionally, the interleavers π_1 are employed for interleaving the BL x_0 , before its XOR-based implantation into the EL x_1 . Following the IL-FEC encoding procedure, the resultant bits are modulated by the binary phase-shift keying (BPSK) modulator of Fig. 1 and then transmitted over *non-dispersive uncorrelated Rayleigh fading wireless channels*.

B. Receiver Model

In this section, we exemplify the IL decoding process using the BL L_0 and the EL L_1 . At the receiver,³ BPSK demodulation is performed, which generates the LLRs containing the systematic information y_0 , y_{01} and the parity information $y_{0,p}$ and $y_{1,p}$, for the layers L_0 and L_1 , respectively. Following the demodulator, the IL-FEC decoder of Fig. 1 is invoked for exchanging extrinsic information across the two layers. Firstly, the FEC decoder 0 will decode the received information y_0 and $y_{0,p}$ for estimating the LLRs of the bits x_0 of the BL L_0 . Then, the resultant extrinsic LLR information of the BL L_0 will be input to the “VND1-VND3-CND-VND2” block of Fig. 1 for extracting the *a-priori* LLRs $L_a(x_1^i)$ ⁴ of the EL L_1 , which is carried out by following the processing of the LLRs in the VND 1, VND 3, CND and VND 2 components of Fig. 2. Specifically, the “VND1-VND3-CND-VND2” block of Fig. 1 performs the following operations step-by-step:

- 1) VND 1 generates the information of BL L_0 for VND 3. The extrinsic LLR $L_e(x_0^i)$ generated by the FEC decoder 0 is input to the VND 1 block of Fig. 1, which extracts the extrinsic LLR information $L_e(x_0^i)$ and forwards it to the VND 3 block of Fig. 1. Since VND 1 has two input branches, it simply duplicates the soft information $L_e(x_0^i)$.
- 2) VND 3 generates the information of BL L_0 for CND. Generally, the summation of the soft information y_0 and the extrinsic LLR $L_e(x_0^i)$ from VND 1 will be output to CND.
- 3) CND generates the information of layer L_1 for VND 2. The LLR information $L_e(x_0^i)$ and the received check information y_{01} is input to the CND block of Fig. 1 for extracting the LLR information of the systematic bit x_1^i , namely the soft input $L_a(x_1^i)$ of VND 2.

³The deinterleavers π^{-1} is not shown in the receiver of Fig. 1 for the sake of simplifying the system architecture.

⁴The subscripts “a” and “e” in L_a and L_e stand for the *a priori* information and extrinsic information [40], respectively.

- 4) VND 2 generates the information of EL L_1 for FEC decoder 1. The LLR information $L_a(x_1^i)$ extracted by the CND is input to the VND 2 block of Fig. 1, which extracts the LLR information $L_a(x_1^i)$ input to the FEC decoder 1 of Fig. 1.

Then, the FEC decoder 1 of Fig. 1 will decode the EL L_1 with the aid of the resultant *a-priori* LLR $L_a(x_1^i)$ and of the soft parity information received from the channel, namely $y_{1,p}$ of Fig. 1. Similar to the “VND1-VND3-CND-VND2” process, the “VND2-CND-VND3-VND1” process is capable of generating the soft information for the BL. With the aid of the “VND1-VND3-CND-VND2” process and the “VND2-CND-VND3-VND1” process, the FEC decoder 0 and FEC decoder 1 may exchange extrinsic information with each other, which constitute an iteration, where the iterative IL decoding process continues until the affordable number of iterations is exhausted or the BL L_0 is perfectly recovered.

III. OPTIMIZED FEC CODING RATES

In this section, we focus our attention on the “Code Rate Optimization” block of our scalable video system seen in Fig. 1, where IL-FEC coded UEP is employed for the sake of achieving an improved video quality. Specifically, the “Code Rate Optimization” block of Fig. 1 characterizes the receiver of Fig. 1 and assists the transmitter in finding the specific FEC coding rates for the layers L_0 and L_1 by minimizing the video distortion. Additionally, although having only two layers is assumed in Section II, more complex coding scenarios may be readily considered, as evidenced by the results of Section IV. Moreover, in our system each layer will be mapped to a single packet. Specifically, in the SVC of H.264 coded video scenario, a network abstraction layer unit (NALU) is packed into a single packet, but the different NALUs may carry a different number of bits. Note that in practice several NALUs may be encapsulated into a single packet, depending on the specific requirements of the network.

Below, we will use a classic RSC codec for describing the proposed solution, which aims for minimizing the reconstructed video distortion at the receiver. This is achieved by characterizing the behavior of the receiver seen in Fig. 1. However, the employment of our proposed techniques is not limited to the RSC codec. Firstly, we define the following notations that will aid our analysis:

- $d(L_0)$: the video distortion induced by the corruption of the BL L_0 , which is measured using the peak signal-to-noise ratio (PSNR);
- $d(L_1)$: the PSNR distortion, when the BL L_0 is correct while the EL L_1 is corrupted;
- $|L_i|$: the length of the bitstream of layer L_i , $0 \leq i \leq 1$;
- R : the overall coding rate of the system shown in Fig. 1;
- r_i : the coding rate of layer L_i , $0 \leq i \leq 1$;
- $p(L_0)$: the packet error ratio (PER) of layer L_0 ;
- $p(L_1)$: the PER of layer L_1 , when the layer L_0 is correctly decoded.

According to the IL-FEC decoding process detailed in Section II, both the FEC decoders 0 and 1 of Fig. 1 affect the PER $p(L_0)$, where FEC decoder 1 feeds back *a-priori*

information about BL L_0 to the FEC decoder 0 through the “VND2-CND-VND3-VND1” blocks of Fig. 1. Specifically, the performance of the FEC decoder 0 depends on SNR , r_0 , $|L_0|$, while the performance of the FEC decoder 1 depends on SNR , r_1 , $|L_1|$. Hence $p(L_0)$ depends on the parameters SNR , $|L_0|$, $|L_1|$, r_0 and r_1 , which can be expressed as

$$p(L_0) = f_0(SNR, |L_0|, |L_1|, r_0, r_1). \quad (2)$$

On the other hand, $p(L_1)$ purely depends on the FEC decoder 1, hence it is determined by the parameters SNR , $|L_1|$ and r_1 , which may be expressed as

$$p(L_1) = f_1(SNR, |L_1|, r_1). \quad (3)$$

Given the specific layers L_0 and L_1 , $d(L_i)$ may be calculated experimentally as the PSNR degrades due to the erasure of layer L_i [20]. Then the expected PSNR distortion induced by BL L_0 may be estimated as $p(L_0) \cdot d(L_0)$. Additionally, when BL L_0 is correctly decoded, the expected PSNR distortion induced by EL L_1 may be estimated as $p(L_1) \cdot d(L_1)$. Hence the expected distortion jointly induced by the decoded layers L_0 and L_1 , denoted as $E[d(L_0, L_1)]$, may be estimated as

$$E[d(L_0, L_1)] = p(L_0) \cdot d(L_0) + [1 - p(L_0)] \cdot p(L_1) \cdot d(L_1). \quad (4)$$

Based on the above discussions, the expected distortion $E[d(L_0, L_1)]$ may be calculated by following the dependency chains displayed in Fig. 3, which relies on the parameters SNR , $|L_0|$, $d(L_0)$, r_0 , $|L_1|$, $d(L_1)$ and r_1 . Among these parameters, SNR , $|L_0|$, $|L_1|$ are determined by the specific layers L_0 , L_1 and the specific transmit power, while $d(L_0)$ and $d(L_1)$ are gleaned from experiments [20]. Hence the coding rates r_0 , r_1 determine the distortion $E[d(L_0, L_1)]$ expected for the layers L_0 , L_1 . Therefore, the objective of the “Code Rate Optimization” block seen in Fig. 1 is to find the specific rates r_0 and r_1 capable of minimizing the expected distortion $E[d(L_0, L_1)]$ of Eq. (4), which may be expressed as

$$\arg \min_{r_0, r_1} \{E[d(L_0, L_1)]\}, \quad (5)$$

subject to the condition of

$$\frac{|L_0|}{r_0} + \frac{|L_1|}{r_1} = \frac{|L_0| + |L_1|}{R}. \quad (6)$$

As illustrated in Fig. 3, Eq. (4) relies on the PER definitions of Eqs. (2) and (3), which cannot be theoretically solved. Below we propose our solution of Eqs. (2) and (3) with the assistance of LUTs and the MI, where LUTs are employed for characterizing the system components of Fig. 1 and MI is employed for numerically quantifying the reliability of the LLRs. Specifically, Section III-A introduces the preliminaries of solving Eqs. (2) and (3), including the LUTs, the MI flow of the VND and CND, followed by the PER estimation of the BL and EL in Section III-B. Then, we exploit both the estimated PERs $p(\cdot)$ and the video distortions $d(\cdot)$ for determining the optimized coding rates in Section III-C. Finally, Section III-D discusses the overheads imposed by the proposed techniques.

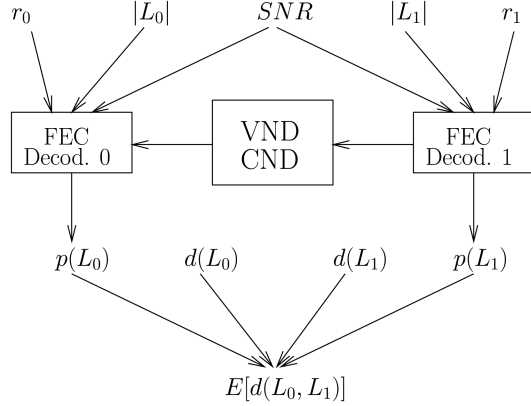


Fig. 3. Dependency chains for calculating the expected distortion formulated in Eq. (4).

A. Preliminaries

The receiver of Fig. 1 consists of the following components: demodulator, FEC decoder, VND, CND. The characteristics of these components jointly determine the PER $p(L_0)$ and $p(L_1)$, as shown in Fig. 3. However, neither the demodulator and nor the FEC decoder may be readily characterized theoretically for diverse system configurations, such as different transceivers, FEC generator polynomials, decoding metrics etc. In order to propose a more universal solution, below LUTs are employed for the sake of characterizing both the demodulator and the FEC decoder, while the MI flow of the VND and CND will be detailed afterwards.

1) *Lookup Tables*:: In our analysis we consider a RSC-coded BPSK-modulated system for transmissions over a *non-dispersive uncorrelated Rayleigh fading wireless channel*. Specifically, we model BPSK based transmissions over a Rayleigh fading wireless channel as a function of the channel SNR and generate the demodulator's output LLRs, where LLRs are quantified by the MI value, as in [41]. Furthermore, *since the MI invoked for quantifying the reliability of the soft information can be applied for Gaussian distributed LLRs [41] generated by arbitrary transceivers, our procedure may be deemed generically applicable, provided that the transceivers generate near-Gaussian distributed LLRs.*⁵ By contrast, the RSC codec is modeled as a function of the SNR, of the LLRs of the systematic bits and of the coding rate, while generating the RSC decoder's output extrinsic LLRs and estimating the PER of the RSC decoded packets, where the LLRs are quantified by the MI value. The following LUTs are created for modeling the transceiver and the RSC codec:

- $T_b(SNR)$: The MI value of the LLRs output by the BPSK demodulator for variable channel SNRs, as exemplified in Table IIa. Since this table relies on the SNR only, it may be stored in a one-dimensional space, where the LUTs' size requirements will be detailed in Section III-D.

- $T_e(SNR, I_s, r)$: The MI value of the extrinsic LLR output of the RSC decoder recorded for variable channel SNRs, as exemplified in Table IIb, where I_s represents the MI value between the *a-priori* LLRs of the systematic information and the corresponding information bits, while r represents the coding rate of the RSC codec. Since this table relies on three parameters, it may be visualized in a three-dimensional space, where the LUTs' size requirements will also be detailed in Section III-D.
- $T_p(SNR, I_s, r)$: The PER value associated with the LLRs output by the RSC decoder at diverse channel SNRs, as exemplified in Table IIb, where I_s represents the MI between the *a-priori* LLRs of the systematic information and the corresponding information bits, while r represents the coding rate of the RSC codec. Note that this LUT may be generated simultaneously with the LUT $T_e(SNR, I_s, r)$, since they have the same input variables.

2) *MI Flow of the VND*:: As seen in Fig. 2a, the operation of the VND may be characterized as $L_{o3}(u_1) = L_{i1}(u_1) + L_{i2}(u_1)$. Assuming that the inputs $L_{i1}(u_1)$ and $L_{i2}(u_1)$ of the VND have the MI values of $I_{i1}[u_1; L_{i1}(u_1)]$ and $I_{i2}[u_1; L_{i2}(u_1)]$ respectively, the MI value of the output $L_{o3}(u_1)$ may be expressed as [42]

$$I_{o3}[u_1; L_{o3}(u_1)] = J \left(\sqrt{J^{-1}(I_{i1}[u_1; L_{i1}(u_1)])^2 + J^{-1}(I_{i2}[u_1; L_{i2}(u_1)])^2} \right), \quad (7)$$

where $J(\cdot)$ and $J^{-1}(\cdot)$ are defined in [37].

3) *MI Flow of the CND*:: Below, we firstly discuss the MI flow of the CND of Fig. 2b, which is designed for the layers L_0 and L_1 , both carrying the same number of bits. Then we will discuss the practical CND employed for realistic unequal length layers.

As seen in Fig. 2b, the action of the CND can be expressed as $L(u_3 = u_1 \oplus u_2) = L(u_1) \boxplus L(u_2)$ [38], where u_1, u_2 and $u_3 = u_1 \oplus u_2$ are random binary variables. Assuming that the inputs $L_i(u_1), L_i(u_2)$ of the CND have the MI values of $I_i[u_1; L(u_1)]$ and $I_i[u_2; L(u_2)]$ respectively, the MI value $I_o[u_3; L(u_3)]$ of u_3 has not been theoretically derived in the literature and since this is not the focus of our work, we model $I_o[u_3; L(u_3)]$ as

$$I_o[u_3; L(u_3)] = T_I(I_i[u_1; L(u_1)], I_i[u_2; L(u_2)]), \quad (8)$$

which is also characterized using LUT.

Given the MI $I_o[u_3; L(u_3)]$ in Eq. (8), we now discuss the practical scenario of the CND, where L_0 and L_1 may contain an unequal number of bits. As detailed in [31], the encoding architecture of L_0 and L_1 may be categorized into two scenarios, namely $|L_0| > |L_1|$ and $|L_1| > |L_0|$. Below we consider the scenario of generating $L(x_1^i)$, when $|L_0| > |L_1|$, which can be generalized to other scenarios. As detailed in [31], the bit x_{01}^i may be generated as

$$x_{01}^i = x_1^i \oplus \sum_{j=1}^n \oplus x_0^j, \quad (9)$$

⁵Our experiments, not included here for reasons of space-economy, suggest that higher order Quadrature Amplitude Modulation (QAM) can be readily applied for transmission over Rayleigh fading wireless channels, since it generates near-Gaussian distributed LLRs.

where \oplus denotes the binary XOR operation. Additionally, since on average⁶ $|L_0|$ bits of L_0 were allocated to the $|L_1|$ bits of L_1 for generating the bits x_{01} , we have $n \in \left\{ \left\lceil \frac{|L_0|}{|L_1|} \right\rceil, \left\lceil \frac{|L_0|}{|L_1|} \right\rceil - 1 \right\}$ [31] for Eq. (9). More specifically, $(N - \frac{|L_0|}{|L_1|})$ proportion of the bits in the EL L_1 are encoded based on Eq. (9) using $n = N - 1$, while the remaining bits of the EL L_1 are encoded based on Eq. (9) relying on $n = N$, where we have $N = \left\lceil \frac{|L_0|}{|L_1|} \right\rceil$.

Based on Eq. (9), the corresponding LLRs $L(x_1)$ of the EL L_1 may be expressed as

$$L(x_1^i) = L(x_{01}^i) \oplus \sum_{j=1}^n \oplus L(x_0^j). \quad (10)$$

Then, the LLRs of $L(x_1)$ are constituted by a mixture of LLRs $L^N(x_1)$ and $L^{N-1}(x_1)$, which are calculated based on Eq. (10) in conjunction with $n = N$ and $n = N - 1$, respectively. We denote the MI values corresponding to the LLRs $L^N(x_1)$ and $L^{N-1}(x_1)$ as $I^N[x_1; L(x_1)]$ and $I^{N-1}[x_1; L(x_1)]$. Based on Eq. (8), the MI values $I^N[x_1; L(x_1)]$ and $I^{N-1}[x_1; L(x_1)]$ may be recursively calculated as in Eq. (11), where $1 \leq k \leq N$. Empirically, we assume that the LLRs $L^N(x_1)$ and $L^{N-1}(x_1)$ are near-Gaussian distributed, while their variance may be denoted as $J^{-1}(I^N[x_1; L(x_1)])$ and $J^{-1}(I^{N-1}[x_1; L(x_1)])$, respectively [37]. Again we assume that $L(x_1)$ exhibits near-Gaussian distribution, hence the variance of $L(x_1)$ may be formulated as shown in (11) and (12) at the bottom of the page.

Finally, the MI value of $L(x_1)$ may be expressed using the J function as in Eq. (13), shown at the bottom of the page.

⁶On average, each bit of L_1 is encoded by $\frac{|L_0|}{|L_1|}$ bits of L_0 for generating a single bit of x_{01} , as detailed in [31].

B. PER Estimation

The PER of the decoded BL and EL having coding rates of r_0 and r_1 will be estimated in Section III-B.1 and Section III-B.2, respectively.

1) *IL-FEC Coded BL*: As shown in Fig. 3, the PER of the BL L_0 depends on the characteristics of the IL-FEC decoder, including the generator polynomials, the trellis decoding techniques employed, as well as the VND and the CND of Fig. 2. Below we propose our solution conceived for estimating the PER of Eq. (3) with the aid of the LUTs defined in Section III-A.1.

The PER estimation flow-chart of the IL-FEC coded BL is illustrated in Fig. 4, which follows the decoding process detailed in Section II. Given the SNR , the MI value I_s of y_0 can be generated by the LUT $T_b(SNR)$. Furthermore, the MI value of the extrinsic LLR gleaned from the RSC decoder 0 may be expressed as $T_e[SNR, T_b(SNR), r_0]$. Then the “VND1-VND3-CND-VND2” process of Fig. 4 is capable of calculating the MI value $I_a(x_1)$ of $L_a(x_1^i)$. Afterwards, based on the LUT $T_e(SNR, I_s, r)$ of the RSC decoder 1, the MI value $I_e(x_1)$ of the extrinsic information may be readily expressed as

$$I_e(x_1) = T_e[SNR, T_b(SNR), r_1]. \quad (14)$$

Then, following the “VND2-CND-VND3-VND1” process of Fig. 4, the improved MI value $I_a(x_0)$ of $L_a(x_0^i)$ can be calculated. Finally, the PER associated with the packet length of l ,⁷ denoted as $p(l)$, may be estimated as $T_p[SNR, I_a(x_0), r_0]$, which is also shown in Fig. 4. Below, we now detail the

⁷The packet containing l bits is employed for generating the LUT $T_p[SNR, I_a(x_0), r_0]$.

$$I^k[x_1; L(x_1)] = \begin{cases} T_I(I[x_0; L(x_0)], I[x_{01}; L(x_{01})]), & k = 1, \\ T_I(I^{k-1}[x_1; L(x_1)], T_I(I[x_0; L(x_0)], I[x_{01}; L(x_{01})])), & k > 1. \end{cases} \quad (11)$$

$$\begin{aligned} \sigma^2 &= D(L(x_1)) \\ &= \left(1 - N + \frac{|L_0|}{|L_1|}\right) \cdot D(L^N(x_1)) \\ &\quad + \left(N - \frac{|L_0|}{|L_1|}\right) \cdot D(L^{N-1}(x_1)) \\ &= \left(1 - N + \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^N[x_1; L(x_1)])^2 \\ &\quad + \left(N - \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^{N-1}[x_1; L(x_1)])^2. \end{aligned} \quad (12)$$

$$I[x_1; L(x_1)] = J(\sigma) = J\left(\sqrt{\left(1 - N + \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^N[x_1; L(x_1)])^2 + \left(N - \frac{|L_0|}{|L_1|}\right) \cdot J^{-1}(I^{N-1}[x_1; L(x_1)])^2}\right). \quad (13)$$

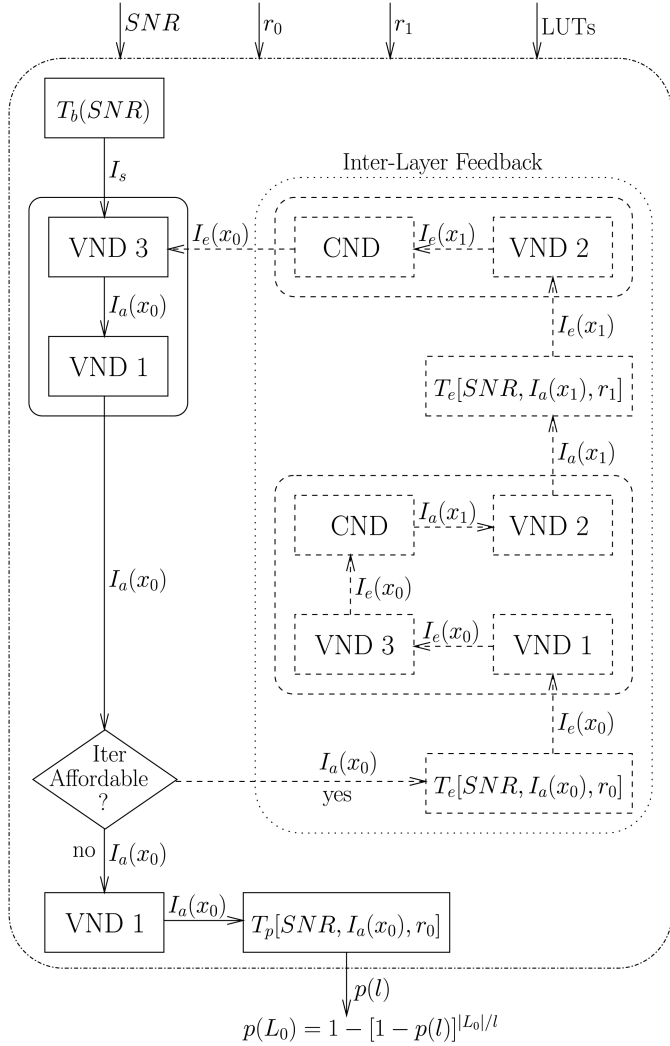


Fig. 4. Flow-chart of the PER estimation for the IL-FEC coded BL L_0 , where $p(L_0)$ will be employed by the “Code Rate Optimization” block of Fig. 1 for resolving the objective function of Eq. (5).

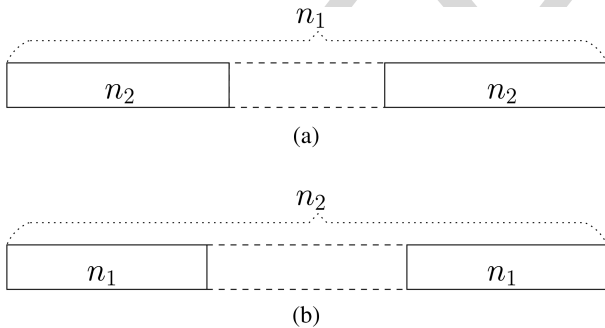


Fig. 5. A packet having a length of $(n_1 \times n_2)$ bits may be divided into n_1/n_2 shorter packets with a length of n_2/n_1 bits. (a) $(n_1 \times n_2)$. (b) $(n_2 \times n_1)$.

method of deriving the PER $p(L_0)$ of the BL L_0 from $p(l) = T_p[SNR, I_a(x_0), r_0]$.

The burst error distribution of RSC codec has been investigated in [43], which is independent of the packet length. Let us consider a RSC-decoded packet having a length of $(n_1 \times n_2)$ bits. Then this packet may be partitioned in two ways, as displayed in Fig. 5. Specifically, it may be divided into n_1 packets,

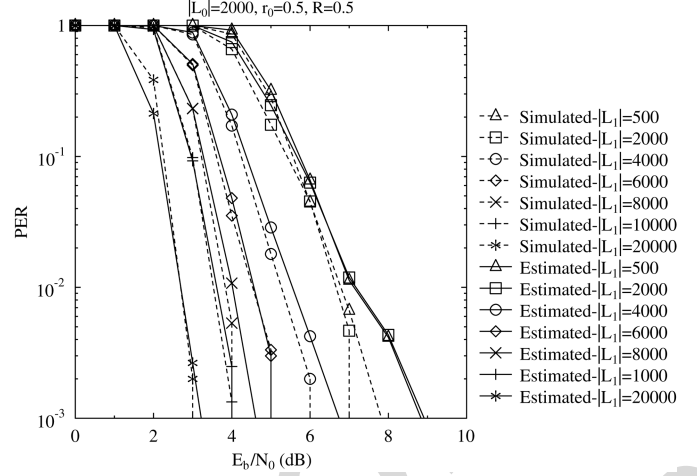


Fig. 6. Comparison of the simulated and estimated PER of the BL, where the parameters $|L_0| = 2000$, $r_0 = 0.5$, $R = 0.5$ are employed.

each carrying n_2 bits or n_2 packets associated with n_1 bits each. Assuming that $p(n_i)$ indicates the PER of the n_i -bit packet, based on Fig. 5a the PER $p(n_1 \cdot n_2)$ may be estimated as

$$p(n_1 \cdot n_2) = 1 - [1 - p(n_2)]^{n_1}, \quad (15)$$

where $p(n_2)$ is the PER of the n_1 packets of Fig. 5a. Similarly, we have $p(n_1 \cdot n_2) = 1 - [1 - p(n_1)]^{n_2}$ based on Fig. 5b. Then, for arbitrary numerical values of n_1, n_2 we have

$$p(n_1) = 1 - [1 - p(n_2)]^{n_1/n_2}. \quad (16)$$

Upon assuming that n_1, n_2 of Eq. (16) are given by $|L_0|$ and l , respectively, the PER $p(L_0)$ of the BL L_0 may be estimated as

$$p(L_0) = f_0(SNR, |L_0|, |L_1|, r_0, r_1) = 1 - [1 - T_p[SNR, I_a(x_0), r_0]]^{|L_0|/l}, \quad (17)$$

where l is the packet length employed for generating the LUT T_p .

For visualizing the accuracy of our proposed PER estimation for the BL, we assume $|L_0| = 2000$, $r_0 = 0.5$ and $R = 0.5$, while the RSC codec is configured by the generator polynomials of [1011, 1101, 1111]. Diverse EL packet length of L_1 are tested. The resultant simulated and estimated PERs of the BL L_0 are displayed in Fig. 6 for different SNR values. As observed in Fig. 6, the estimated PER does not perfectly match the simulated PER, but it closely follows the same trend, especially in the PER region above 10^{-2} , which is our area of interest. Note that the estimation error may be introduced both by the LUTs and by the near-Gaussian distributed LLRs output by the demodulator.

2) *FEC Coded EL*:: When the BL L_0 is correctly decoded, infinite LLRs will be input to the CND 3 of Fig. 1 [31], hence the LLRs having the MI value of I_s will be input to the VND 2 by the CND. Therefore the PER of the EL having a length of C bits may be estimated as $T_p(SNR, I_s, r_1)$ with the aid of the LUT T_p . Then, similar to the Eq. (17), the PER $p(L_1)$ of the EL L_1 may be finally estimated as

$$p(L_1) = f_1(SNR, |L_1|, r_1) = 1 - [1 - T_p(SNR, I_s, r_1)]^{|L_1|/l}, \quad (18)$$

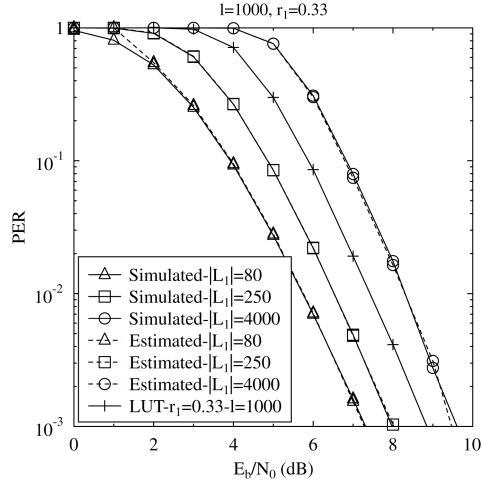


Fig. 7. Comparison of the simulated and estimated PER for the EL, where the parameters $l = 1000$, $r_1 = 0.33$ are employed.

where l is the packet length employed for generating the LUT T_p .

For visualizing the accuracy of our proposed EL PER estimation, we assume $r_1 = 0.33$, while the RSC codec is configured by the generator polynomials of $[1011, 1101, 1111]$. Moreover, we consider $|L_1|$ values ranging from 80 to 4000, while the LUT $T_p(SNR, I_s, r_1)$ is generated based on $l = 1000$. The resultant simulated and estimated PER of the EL L_1 is displayed in Fig. 7. Observe in Fig. 7 that the estimation error is much smaller than that in Fig. 6 due to the fact that the EL is protected by an RSC codec without using IL techniques.

C. Optimized Coding Rates

Based on Section III-B, the PERs of L_0 and L_1 are estimated according to the specified coding rates of r_0 and r_1 for the BL and EL, respectively. Hence the expected distortion $E[d(L_0, L_1)]$ imposed by the wireless transmission using rates of r_0 and r_1 may be readily calculated using Eq. (4). In order to find the rates r_0, r_1 minimizing the distortion $E[d(L_0, L_1)]$ in Eq. (5), a full search of the coding rate space may be performed. Note that r_1 is determined by Eq. (6) for a given r_0 , hence the full search is performed in a one-dimensional space to find the optimal r_0 .

To elaborate a little further, we consider the scenario of $SNR = 10$ dB, $R = 1/3$, $|L_0| = |L_1| = 100$, $d(L_0) = 20$, $d(L_1) = 10$. Based on the conditions formulated in Eq. (6), the coding rate r_1 of L_1 is determined as follows:

$$r_1 = \frac{|L_1|}{\frac{|L_0| + |L_1|}{R} - \frac{|L_0|}{r_0}} = \frac{r_0}{6 \cdot r_0 - 1} \quad (19)$$

$$r_0 \leq 1, r_1 \leq 1.$$

Then, based on Eqs. (2), (3) and (4) the expected value of $E[d(L_0, L_1)]$ in Eq. (5) may be formulated as

$$\begin{aligned} E[d(L_0, L_1)] &= p(L_0) \cdot d(L_0) + [1 - p(L_0)] \cdot p(L_1) \cdot d(L_1) \\ &= 20 \cdot f_0 \left(10, 100, 100, r_0, \frac{r_0}{6 \cdot r_0 - 1} \right) \end{aligned}$$

$$\begin{aligned} &+ 10 \cdot \left[1 - f_0 \left(10, 100, 100, r_0, \frac{r_0}{6 \cdot r_0 - 1} \right) \right] \\ &\cdot f_1 \left(10, 100, \frac{r_0}{6 \cdot r_0 - 1} \right) \\ &= g(r_0), \end{aligned} \quad (20)$$

where the functions f_0 and f_1 are discussed in the ‘PER Estimation’ Section III-B. Note that we introduced the short-hand of $E[d(L_0, L_1)] = g(r_0)$ in Eq. (20) for simplicity. Based on Eq. (19), we can readily arrive at $\frac{1}{5} \leq r_0 \leq 1$. Then we artificially limit r_0 to the following discrete practical values $[0.2 : 0.05 : 1]$.⁸ Finally, the solution r_0 of the objective function $\arg \min \{E[d(L_0, L_1)]\}$ in Eq. (5) may be readily found to be:

$$\left\{ r_0 \mid \arg \min_{r_0} g(r_0), r_0 \in \{0.2 : 0.05 : 1\} \right\}, \quad (21)$$

which can be solved by performing a one-dimensional full search. Then, given r_0 the coding rate r_1 may be readily expressed from Eq. (19).

D. Overheads

All the optimization operations detailed in Section III are carried out at the transmitter. Below, we discuss the overheads imposed at the transmitter, while the overheads imposed at the receiver are given in [31]. The overheads imposed at the transmitter include the estimation of $d(\cdot)$, the generation of LUTs, the estimation of $p(\cdot)$, the imposed frame delay and the full search detailed in Section III-C. Among these overheads, the generation of LUTs only imposes extra off-line design-time, while the estimation of $d(\cdot)$, $p(\cdot)$ and the full search impose extra on-line run-time complexity. For notational simplicity, we use n_{snr}, n_I, n_r to denote the number of the parameters SNR, I_s, r , respectively, which are used for generating the LUTs. For example, if the values of I_s range from 0 to 1 with a step size of 0.01, we have $n_I = 101$. In the following, we will analyze these overheads in order to validate our system.

1) *Estimation of $d(\cdot)$* : $d(L_i)$ is estimated in a similar manner to the procedure of [20], where the distortion $d(L_0)$ may be obtained by decoding the bitstream in the presence of a corrupted BL L_0 . Alternatively, the solutions of [21], [25], [44] may be applied in our system.

2) *Generation of LUTs*: Three LUTs were generated in our proposed solution. The LUT $T_b(SNR)$ characterizes the channel and transceiver, hence this table has to be regenerated when the channel or the transceiver are changed. However, this LUT only has a single variable, hence it is straight-forward to generate the LUT $T_b(SNR)$. Additionally, the LUTs $T_e(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$ can be simultaneously generated by simulations, since they have the same variables. However, they are independent of the channel and transceiver, but they are dependent on the configuration of the FEC code. Hence the LUTs T_e and T_p have to be regenerated, when the FEC is reconfigured. *Note that all the LUTs are independent of the video sequences and that these LUTs are generated during the design process.* The size of these LUTs depends on

⁸The value of r_0 ranges from 0.2 to 1 with a constant step of 0.05.

the variables, hence the LUTs T_b , T_e , T_p have sizes of n_{snr} , $(n_{snr} \cdot n_I \cdot n_r)$, $(n_{snr} \cdot n_I \cdot n_r)$, respectively.

3) *Estimation of $p(\cdot)$* : The CND and VND are involved in the PER estimation process, as well as in the search through the LUTs. As detailed in Section III-A, the MI flows of the VND and CND impose a low computational complexity. Furthermore, in our system the values of each variable are chosen using a constant step-size, which guarantees low complexity, while searching through the LUTs. For example, let us assume that the vector $[0 : 0.01 : 1]$ is employed for representing the values of I_s for generating the LUTs. In this case, the I_s value of 0.5 can be directly located at index 50.

4) *Full Search of r_0* : n_r coding rates of r_0 may be tentatively tested, while r_1 is determined by Eq. (6) for a given r_0 . Hence the complexity of the full search may be expressed as $O(C \times n_r)$, where C is a constant, representing the complexity of estimating $p(L_i)$. However, when more layers are encapsulated in the source bitstream, the full-search based complexity increases exponentially, leading to a multi-dimensional optimization problem, which has been widely studied in the literature [45]–[47]. Specifically, the adaptive particle swarm optimization (APSO) technique of [47] may be readily employed for finding the global optimum in real-time. Note that in the scenarios where as few as 2–4 layers in the range of 2 to 4 are generated, even elite-force full-search is realistic at a modest complexity.

5) *Delay*: Our technique is implemented using the SVC of H.264 as detailed in Section IV, where each video frame can be encoded into a single BL and multiple ELs. Since the IL encoding and decoding process is performed within each frame, no extra delay is imposed by our proposed technique. Hence, the system can be applied in low-delay applications.

IV. SYSTEM PERFORMANCE

In this section, we benchmark our proposed Optimized-IL-RSC system against an identical RSC-aided system employing the traditional UEP technique. The generator polynomials of [1011, 1101, 1111] are employed for configuring the RSC codec. Moreover, BPSK modulated signals were transmitted through *non-dispersive uncorrelated Rayleigh fading wireless channels*. The 30-frame Foreman and Football video clips represented in (176×144) -pixel quarter common intermediate format (QCIF) and 4 : 2 : 0 YUV format were encoded by the JSVM H.264/AVC reference video codec. The Football and Foreman sequences exhibit different motion-activity, hence they allow us to demonstrate the universal nature of our proposed system. These sequences were scanned at 15 and 30 FPS, respectively. The “frame-copy” based error concealment tool built into the JSVM H.264/AVC reference codec was activated for combating the effects of channel impairments. The GOP duration was set to 15, hence an Intra-frame-coded (I) picture was inserted every 15 frames. Correspondingly, both of the two video sequences were encoded into GOPs, consisting of an I frame, followed by 14 predicted (P) frames. Since the bi-directionally predicted (B) frames may impose error propagation on their forward- and backward-predicted dependent-frames, the B frames are disabled in our JSVM configuration. Additionally, only the Medium Grain Scalability (MGS) [4], [48] feature is

TABLE I
THE PARAMETERS OF THE VIDEO SEQUENCES EMPLOYED

	Football	Foreman
Representation	YUV 4:2:0	YUV 4:2:0
Format	QCIF	QCIF
Bits Per Pixel	8	8
FPS	15	30
Number of Frames	30	30
Video Codec	SVC-H.264	SVC-H.264
Bitrate	2297 kbps	218 kbps
Error-Free PSNR	38.8 dB	37.3 dB
Error Concealment	Frame-Copy	Frame-Copy

enabled, when encoding the video sequences into three different-quality ELs, namely into the layers L_0 , L_1 and L_2 using the standardized quantization parameters (QP) of 40, 32 and 24, respectively. Furthermore, each frame will be encoded into a single BL and two ELs, resulting in three network abstraction layer units (NALUs). These configurations jointly result in a bitrate of 2297 kbps and a PSNR of 38.8 dB for the Football sequence in the absence of transmission errors, while achieving 37.3 dB PSNR at 218 kbps for the Foreman sequence. The system parameters are listed in Table I.

Moreover, each SVC H.264/AVC-compressed bitstream was channel encoded and transmitted on a NALU by NALU [5] basis, which is the smallest unit to be decoded by the SVC decoder. Each NALU was protected by CRC codes. At the receiver, each decoded NALU failing to pass the CRC check process was removed before the SVC video decoding process. In all of our experiments, the compressed bitstreams were transmitted 300 times in order to generate statistically sound performance curves.

A. Off-line LUTs Generation

In our experiments, the vectors of $[0 : 0.5 : 15]$, $[0 : 0.01 : 1]$, $[0.33 : 0.02 : 1]$ ⁹ are utilized for the variables SNR , I_s , r , respectively, for generating the LUTs, which result in $n_{snr} = 31$, $n_I = 101$, $n_r = 33$. Firstly, we recorded the MI values of the LLRs output by the BPSK demodulator for the SNR values of $[0 : 0.5 : 15]$ for the sake of generating the LUT $T_b(SNR)$. Then for each SNR , the MI values $[0 : 0.01 : 1]$ for I_s and the coding rates $[0.33 : 0.02 : 1]$ for r were utilized for generating the LUTs $T_e(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$, simultaneously. Furthermore, 8-byte floating values were utilized for storing the LUTs in memory. Correspondingly, the LUTs T_b , T_e and T_p require memory sizes of about 248 bytes, 800 KB and 800 KB, respectively. Finally, $n_r = 33$ results in a full search complexity of $(C \times n_r^2 = C \cdot 1089)$, thereby finding the optimized coding

⁹These values can be stored as floats in 8 bytes each.

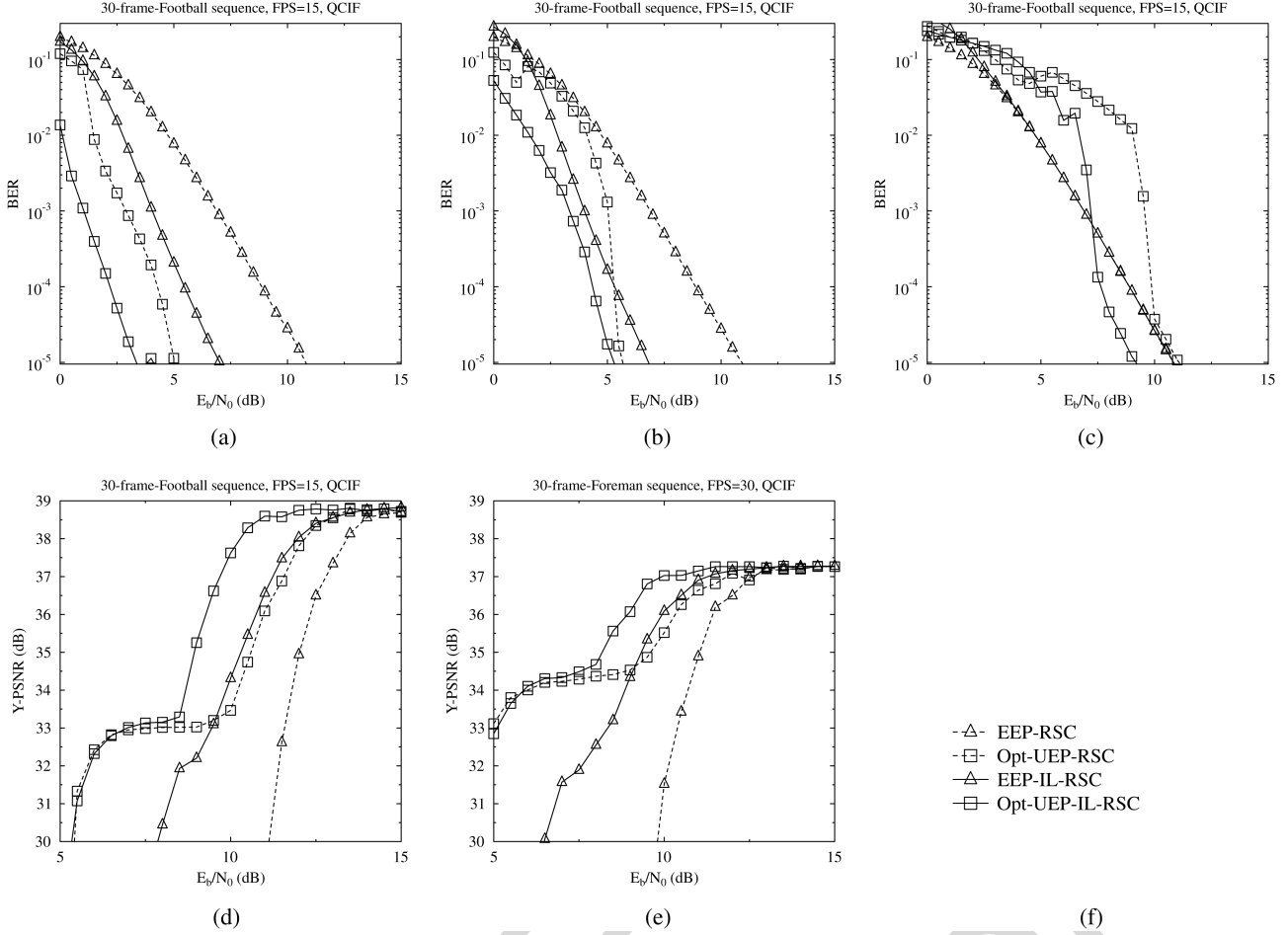


Fig. 8. BER and PSNR vs E_b/N_0 performance comparison of the proposed system and the benchmarks, namely the EEP-RSC scheme, the Opt-UEP-RSC scheme [20], the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme for the *Football* and *Foreman* sequences. *Non-dispersive uncorrelated Rayleigh fading wireless channels*. (a) BER vs E_b/N_0 for L_0 -*Football*. (b) BER vs E_b/N_0 for L_1 -*Football*. (c) BER vs E_b/N_0 for L_2 -*Football*. (d) PSNR vs E_b/N_0 for *Football*. (e) PSNR vs E_b/N_0 for *Foreman*. (f) Tested schemes.

TABLE II
EXAMPLE OF THE LUTs (a) $T_b(SNR)$.
(b) $T_c(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$

SNR	I_s	SNR	I_a	r	I_e	$p(l)$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
1	0.45	3	0.67	0.87	0.81	0.93
2	0.50	3	0.67	0.88	0.83	0.89
3	0.56	3	0.67	0.89	0.85	0.84
4	0.62	3	0.67	0.90	0.87	0.77
5	0.67	3	0.67	0.91	0.89	0.69
6	0.72	3	0.67	0.92	0.91	0.57
7	0.76	3	0.67	0.93	0.93	0.46
8	0.80	3	0.67	0.94	0.94	0.33
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

(a) $T_b(SNR)$ (b) $T_c(SNR, I_s, r)$ and $T_p(SNR, I_s, r)$

rates of r_0 , r_1 and r_2 . A number of entries of the LUTs generated for our system are displayed in Table II.

B. System Performance

In this section, we benchmark our system against the traditional equal error protection (EEP) system, which is referred to

as EEP-RSC. Furthermore, the traditional optimal UEP-RSC (Opt-UEP-RSC) system is also presented, which is the system of [20] applied in the scenario of an RSC coded system, where all the layers are unequally protected by an RSC code dispensing with the IL technique. Moreover, the IL aided EEP system (EEP-IL-RSC) is also considered.

The BER curves of the L_0 , L_1 and L_2 layers of the *Football* sequence are displayed in Figs. 8a, 8b and 8c, respectively, where the Opt-UEP-IL-RSC system always outperforms the Opt-UEP-RSC system. Moreover, as seen in Fig. 8c, the EEP schemes outperform the Opt-UEP-IL-RSC and Opt-UEP-RSC schemes in the lower E_b/N_0 range. This is due to the fact that the coding-rate of layer L_2 is sacrificed for the sake of protecting the L_0 and L_1 layers. Similar trends were observed for the *Foreman* sequence, which are not included here owing to space-economy.

The PSNR versus E_b/N_0 performance recorded for the *Football* sequence is displayed in Fig. 8d, where we observe that the EEP-IL-RSC scheme substantially outperforms the EEP-RSC system, while it only slightly outperforms the Opt-UEP-RSC scheme. However, by optimizing the coding rates, the Opt-UEP-IL-RSC scheme becomes capable of substantially outperforming the Opt-UEP-RSC system. In the low E_b/N_0 range, the Opt-UEP-IL-RSC and Opt-UEP-RSC



Fig. 9. Comparison of frames at E_b/N_0 of 10 dB for the *Football* sequences. The five columns (from left to right) indicate frames of the original video, the EEP-RSC scheme, the Opt-UEP-RSC [20] scheme, the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme, respectively.

schemes have a similar PSNR performance and tend to exhibit a residual error floor. This is due to the fact that these two systems endeavor to protect the BL L_0 by sacrificing the protection of the ELs, where the ELs are unlikely to be recovered before reaching $E_b/N_0 = 10$ dB. This error floor becomes explicit in the BER curves of L_2 displayed in Fig. 8c. On the other hand, the Opt-UEP-RSC scheme outperforms the EEP-RSC scheme by about $E_b/N_0 = 1.3$ dB at a PSNR of 38 dB. The Opt-UEP-IL-RSC scheme achieves an E_b/N_0 reduction of about 1.9 dB compared to the Opt-UEP-RSC scheme at a PSNR of 38 dB. Alternatively, about 3.3 dB of PSNR video quality improvement is observed at an E_b/N_0 of 10 dB. A subjective comparison of the benchmarks recorded for the *Football* sequence is presented in Fig. 9.

For providing further insights for video scenes having different motion-activity, the PSNR versus E_b/N_0 performance of the Opt-UEP-IL-RSC is portrayed in Fig. 8e using the *Foreman* sequence. The Opt-UEP-RSC scheme outperforms the EEP-RSC system by about $E_b/N_0 = 0.8$ dB at a PSNR of 37 dB. Moreover, about 1.7 dB of power reduction is achieved by the Opt-UEP-IL-RSC scheme compared to the Opt-UEP-RSC scheme at a PSNR of 37 dB. Alternatively, about 1.6 dB of PSNR video quality improvement may be observed at an E_b/N_0 of 10 dB.

C. Optimized Coding Rates

The optimized coding rates found by our proposed regime for the *Football* sequence are shown in Fig. 10 in comparison to the benchmarks. Specifically, the y axis of Fig. 10 indicates the averaged coding rates of all frames. Furthermore, the coding rates of L_0 and L_2 are presented, while the curves of L_1 are similar to those of the BL L_0 . Observe in Fig. 10 that in the low E_b/N_0 range the coding rates of the BL L_0 of the Opt-UEP-RSC and Opt-UEP-IL-RSC schemes are lower than the overall coding rate of 0.5. The reason of this observation is that the protection of the layers L_1 and L_2 is sacrificed for the sake of protecting the more important BL L_0 . At high E_b/N_0 values, the coding rates of the BL L_0 are increased due to the fact that L_0 is more likely to be corrected at high coding rates, when favorable channel conditions prevail. In comparison to the BL L_0 , the coding rates of L_2 are reduced upon increasing E_b/N_0 .

V. CONCLUSIONS

We proposed an optimization technique for bit-level IL-FEC coded scalable video transmission over wireless channels,

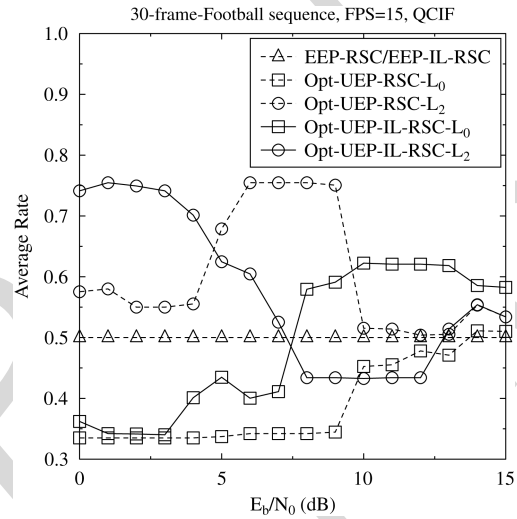


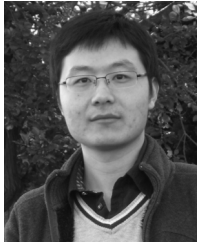
Fig. 10. Coding rates comparison of the benchmarks, namely the EEP-RSC scheme, the Opt-UEP-RSC [20] scheme, the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme, for the *Football* sequence. *Non-dispersive uncorrelated Rayleigh fading wireless channels*

where soft-decoding aided FEC is utilized. Firstly, LUTs were generated for the sake of optimizing the wireless transceivers and the FEC codec, which relied on exploiting the off-line training based MI. Moreover, the MI exchange of the VND and CND of Fig. 2 was exploited for realistic video systems, where both the PERs of the IL-FEC coded BL and of the FEC coded ELs were estimated, based on the knowledgeable specific coding rates of all the scalable video layers. Then the expected PSNR distortion was minimized by performing a full search. BPSK modulation, an RSC FEC code and the SVC-H.264 video codec were employed in our design example, but our solution may be readily applied for arbitrary transceivers generating near-Gaussian distributed LLRs, for arbitrary non-iteratively decoded channel codes and for arbitrary layered video transmissions. Our simulation results show that the proposed optimized IL-FEC system outperforms the traditional optimal UEP system by about 1.9 dB of E_b/N_0 or 3.3 dB of PSNR, when employing a RSC code.

In our future work, we will incorporate the IL-FEC scheme into multiview video coding and we will also carry out further investigations for applying our solution to more advanced FEC codes, such as turbo codes [34], LDPC codes, BICM, BICM-ID, TCM aided TTCM schemes.

REFERENCES

- [1] L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG2, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. New York, NY, USA: Wiley, 2007.
- [2] T. Zhang and Y. Xu, "Unequal packet loss protection for layered video transmission," *IEEE Trans. Broadcast.*, vol. 45, pp. 243–252, Jun. 1999.
- [3] H. Imaizumi and A. Luthra, "Three-Dimensional Television, Video and Display Technologies," in *ch. MPEG-2 Multiview Profile*. Berlin, Germany: Springer Verlag, 2002, pp. 169–181.
- [4] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, pp. 1103–1120, Sep. 2007.
- [5] "Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG," ITU-T Rec. H.264/ISO/IEC 14496-10 AVC: Advanced Video Coding for Generic Audiovisual Services, Mar. 2010.
- [6] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proc. IEEE*, vol. 99, pp. 626–642, Apr. 2011.
- [7] F. Yang, Q. Zhang, W. Zhu, and Y.-Q. Zhang, "End-to-end TCP-friendly streaming protocol and bit allocation for scalable video over wireless Internet," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 777–790, May 2004.
- [8] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Trans. Inf. Theory*, vol. 13, pp. 600–607, Oct. 1967.
- [9] T. Brüggen and P. Vary, "Unequal error protection by modulation with unequal power allocation," *IEEE Commun. Lett.*, vol. 9, pp. 484–486, Jun. 2005.
- [10] V. Pavlushkov, R. Johannesson, and V. Zyablov, "Unequal error protection for convolutional codes," *IEEE Trans. Inf. Theory*, vol. 52, pp. 700–708, Feb. 2006.
- [11] J. Hagenauer, "Rate-compatible puncture convolutional codes (RCPC) and their application," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, Apr. 1988.
- [12] N. Rahnavard and F. Fekri, "New results on unequal error protection using LDPC codes," *IEEE Commun. Lett.*, vol. 10, pp. 43–45, Jan. 2006.
- [13] V. Kumar and O. Milenkovic, "On unequal error protection LDPC codes based on Plotkin-type constructions," *IEEE Trans. Commun.*, vol. 54, pp. 994–1005, Jun. 2006.
- [14] C. Gong, G. Yue, and X. Wang, "Message-wise unequal error protection based on low-density parity-check codes," *IEEE Trans. Commun.*, vol. 59, pp. 1019–1030, Apr. 2011.
- [15] N. Rahnavard, H. Pishro-Nik, and F. Fekri, "Unequal error protection using partially regular LDPC codes," *IEEE Trans. Commun.*, vol. 55, pp. 387–391, Mar. 2007.
- [16] 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.
- [17] S. X. Ng, J. Y. Chung, and L. Hanzo, "Turbo-detected unequal protection MPEG-4 wireless video telephony using multi-level coding, trellis coded modulation and space-time trellis coding," *IEEE Proc. Commun.*, vol. 152, pp. 1116–1124, Dec. 2005.
- [18] H. Wang, F. Zhai, Y. Eisenberg, and A. Katsaggelos, "Cost-distortion optimized unequal error protection for object-based video communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 1505–1516, Dec. 2005.
- [19] M. Aydinlik and M. Salehi, "Turbo coded modulation for unequal error protection," *IEEE Trans. Commun.*, vol. 56, pp. 555–564, Apr. 2008.
- [20] H. Ha and C. Yim, "Layer-weighted unequal error protection for scalable video coding extension of H.264/AVC," *IEEE Trans. Consumer Electron.*, vol. 54, pp. 736–744, May 2008.
- [21] Y. C. Chang, S. W. Lee, and R. Komiya, "A fast forward error correction allocation algorithm for unequal error protection of video transmission over wireless channels," *IEEE Trans. Consumer Electron.*, vol. 54, pp. 1066–1073, Aug. 2008.
- [22] Y. C. Chang, S. W. Lee, and R. Komiya, "A low complexity hierarchical QAM symbol bits allocation algorithm for unequal error protection of wireless video transmission," *IEEE Trans. Consumer Electron.*, vol. 55, pp. 1089–1097, Aug. 2009.
- [23] D. Sejdinović, D. Vukobratović, A. Doufexi, V. Šenk, and R. Piechocki, "Expanding window fountain codes for unequal error protection," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2510–2516, 2009.
- [24] D. Vukobratović, V. Stanković, D. Sejdinović, L. Stanković, and Z. Xiong, "Scalable video multicast using expanding window fountain codes," *IEEE Trans. Multimedia*, vol. 11, no. 6, pp. 1094–1104, 2009.
- [25] E. Maani and A. Katsaggelos, "Unequal error protection for robust streaming of scalable video over packet lossy networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, pp. 407–416, Mar. 2010.
- [26] S. Ahmad, R. Hamzaoui, and M. Al-Akaidi, "Unequal error protection using fountain codes with applications to video communication," *IEEE Trans. Multimedia*, vol. 13, pp. 92–101, Feb. 2011.
- [27] K. Nguyen, T. Nguyen, and S.-C. Cheung, "Video streaming with network coding," *J. Signal Process. Syst.*, vol. 59, pp. 319–333, Jun. 2010.
- [28] C. Hellge, D. Gomez-Barquero, T. Schierl, and T. Wiegand, "Layer-aware forward error correction for mobile broadcast of layered media," *IEEE Trans. Multimedia*, vol. 13, pp. 551–562, Jun. 2011.
- [29] Nasruminallah and L. Hanzo, "Near-capacity H.264 multimedia communications using iterative joint source-channel decoding," *IEEE Commun. Surveys and Tutorials*, vol. 14, pp. 538–564, 2012, Second Quarter.
- [30] M. Halloush and H. Radha, "Network coding with multi-generation mixing: A generalized framework for practical network coding," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 466–473, 2011.
- [31] Y. Huo, M. El-Hajjar, and L. Hanzo, "Inter-layer FEC aided unequal error protection for multilayer video transmission in mobile TV," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1622–1634, 2013.
- [32] D. Wu, Y. T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the Internet: Challenges and approaches," *Proc. IEEE*, vol. 88, no. 12, pp. 1855–1877, 2000.
- [33] T. Stockhammer, M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657–673, 2003.
- [34] L. Hanzo, T. Liew, B. Yeap, and R. Tee, *Turbo Coding, Turbo Equalization and Space-Time Coding*. New York, NY, USA: Wiley, 2011.
- [35] T. Stockhammer, A. Shokrollahi, M. Watson, M. Luby, and T. Gasiba, "Application Layer Forward Error Correction for Mobile Multimedia Broadcasting," in *Handbook of Mobile Broadcasting: DVB-H, DMB, ISDB-T and Media FLO*, B. Furhet and S. Ahson, Eds. Boca Raton, FL, USA: CRC, 2008, pp. 239–280.
- [36] C. Hellge, T. Schierl, and T. Wiegand, "Multidimensional layered forward error correction using rateless codes," in *Proc. IEEE Int. Conf. Communications*, Beijing, China, May 2008, pp. 480–484.
- [37] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, pp. 670–678, Apr. 2004.
- [38] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, pp. 429–445, Mar. 1996.
- [39] J. Chen, A. Dholakia, E. Eleftheriou, M. Fossorier, and X.-Y. Hu, "Reduced-complexity decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 53, pp. 1288–1299, Aug. 2005.
- [40] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE Int. Conf. Communications*, Geneva, Switzerland, May 1993, pp. 1064–1070.
- [41] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, Oct. 2001.
- [42] R. Maunder and L. Hanzo, "Extrinsic information transfer analysis and design of block-based intermediate codes," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 762–770, Mar. 2011.
- [43] R. Khalili and K. Salamatian, "A new analytic approach to evaluation of packet error rate in wireless networks," in *Proc. Communication Networks and Services Research Conf.*, Halifax, NS, Canada, May 2005, pp. 333–338.
- [44] X. Yang, C. Zhu, Z. Li, X. Lin, G. Feng, S. Wu, and N. Ling, "Unequal loss protection for robust transmission of motion compensated video over the Internet," *Signal Process.: Image Commun.*, vol. 18, no. 3, pp. 157–167, 2003.
- [45] M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Computat.*, vol. 6, no. 1, pp. 58–73, 2002.
- [46] D. Schonfeld and N. Bouaynaya, "A new method for multidimensional optimization and its application in image and video processing," *IEEE Signal Process. Lett.*, vol. 13, no. 8, pp. 485–488, 2006.
- [47] Z.-H. Zhan, J. Zhang, Y. Li, and H.-H. Chung, "Adaptive particle swarm optimization," *IEEE Trans. Syst., Man, Cybern. B: Cybern.*, vol. 39, no. 6, pp. 1362–1381, 2009.
- [48] A. Detti, G. Bianchi, C. Pisa, F. Proto, P. Loreti, W. Kellerer, S. Thakolsri, and J. Widmer, "SVEF: An open-source experimental evaluation framework for H.264 scalable video streaming," in *Proc. IEEE Symp. Computers and Communications, ISCC*, Sousse, Tunisia, Jul. 2009, pp. 36–41.



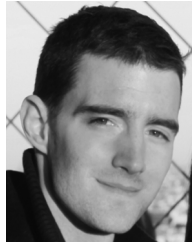
Yongkai Huo received the B.Eng. degree with distinction in computer science and technology from Hefei University of Technology, Hefei, China, in 2006 and the M.Eng. degree in computer software and theory from University of Science and Technology of China, Hefei, China, in 2009. He is currently working toward the Ph.D. degree with the Communications, Signal Processing and Control Group, School of Electronics and Computer Science, University of Southampton, Southampton, U.K. He received a scholarship under the China-U.K. Scholarships for Excellence Programme.

His research interests include distributed video coding, multiview video coding, robust wireless video streaming and joint source-channel decoding.



Mohammed El-Hajjar is a lecturer in the Electronics and Computer Science in the University of Southampton. He received his BEng degree in Electrical Engineering from the American University of Beirut, Lebanon in 2004. He then received an MSc in Radio Frequency Communication Systems and PhD in Wireless Communications both from the University of Southampton, UK in 2005 and 2008, respectively. Following the PhD, he joined Imagination Technologies as a research engineer, where he worked on designing and developing the

BICM peripherals in Imagination's multi-standard communications platform, which resulted in several patent applications. In January 2012, he joined the Electronics and Computer Science in the University of Southampton as a lecturer in the Communications, Signal Processing and Control research group. He is the recipient of several academic awards and has published a Wiley-IEEE book and in excess of 30 journal and international conference papers. His research interests are mainly in the development of intelligent communications systems for the Internet of Things including massive MIMO systems for mm-wave communications, cooperative communications and Radio over fibre systems.



Robert G. Maunder has studied with Electronics and Computer Science, University of Southampton, UK, since October 2000. He was awarded a first class honors BEng in Electronic Engineering in July 2003, as well as a PhD in Wireless Communications and a lectureship in December 2007. His research interests include joint source/channel coding, iterative decoding, irregular coding and modulation techniques. He has published a number of IEEE papers in these areas.



Lajos Hanzo (F'04) received the degree in electronics in 1976, the Doctorate degree in 1983, and the D.Sc. degree.

During his 35-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 80 PhD students, co-authored 20 John Wiley/IEEE Press books on mobile radio com-

munications totalling in excess of 10 000 pages, published 1300 research entries at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European IST Programme and the Mobile Virtual Centre of Excellence (VCE), UK. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and since 2009 he has been a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to <http://www-mobile.ecs.soton.ac.uk>.